

# University of St Andrews



Full metadata for this thesis is available in  
St Andrews Research Repository  
at:

<http://research-repository.st-andrews.ac.uk/>

This thesis is protected by original copyright

# Computer graphic manipulations in the study of face perception

David Michael Burt





TL E190

## ***Declarations:***

I, David Michael Burt, hereby certify that this thesis, which is approximately 57,000 words in length, has been written by me, that it is the record of work carried out by me and that it had not been submitted in any previous application for a higher degree.

Date. 17/9/02 Signature of candidate.

I was admitted as a research student in November 1996 and as a candidate for the degree of PhD in March 1998; the higher study for which this is a record was carried out in the University of St Andrews between 1996 and 2002.

Date. 17/9/02 Signature of candidate.

I hereby certify that the candidate has fulfilled the conditions of the Resolution and Regulations appropriate for the degree of PhD in the University of St Andrews and that the candidate is qualified to submit this thesis in application for that degree.

Date. 17/9/02 Signature of supervisor..

In submitting this thesis to the University of St Andrews I understand that I am giving permission for it to be made available for use in accordance with the regulation of the University Library for the time being in force, subject to any copyright vested in the work not being affected thereby. I also understand that the title and abstract will be published and that a copy of the work may be made and supplied to any bona fide library or research worker.

Date. 17/9/02 Signature of candidate.

***Thesis copyright:***

Unrestricted

In submitting this thesis to the University of St Andrews I understand that I am giving permission for it to be made available for use in accordance with the regulations of the University Library for the time being in force, subject to any copyright vested in the work not being affected thereby. I also understand that the title and abstract will be published, and that a copy of the work may be made and supplied to any bona fide library or research worker.

Date...17/9/02.....signature of candidate,

.....

## ***Acknowledgements***

I am very grateful to a number of people for their help in the production and delivery of this PhD:

Dave Perrett, my supervisor, for his support and for providing an inspirational environment and the perfect balance between giving direction and encouraging independent thought.

Elisa Frigerio, a model of determination, without whom the work on facial expressions presented here would never have been done or thought of.

Tony Little for many discussions about attractiveness and much damage inflicted to my liver.

Bernie Tiddeman and Duncan Rowland both for their help and ideas related to programming and for the use of their own software.

Ben Jones, Keith Payne and Barbara Montague for their work on evolutionary psychology and facial expression perception.

I am neither the best nor the most motivated writer and I am very grateful for the help and encouragement of a number of other people in the preparation of this manuscript. Most especially Stefanie Sonnenberg, Eva Loth, Elisabeth Cornwell, Anne Perrett.

And, to my parents, for their help and encouragement and for making everything (including me) possible.

## Contents

Declarations:	2
Thesis copyright:	3
Acknowledgements	4
Contents	5
Abstract	9
Note on thesis publications	10
<i>Chapter 1: An introduction to computer graphics techniques and their uses in the study of face perception.</i>	<i>11</i>
Introduction	12
Line-drawn computer caricatures	13
Image warping	22
Image blending using feature points	27
Image morphing	40
Manipulation of individual features	51
Image and shape caricaturing	55
Image and shape transforms	59
Texture processing	64
PCA analysis of facial variation	71
Genetic algorithms as a method of searching face space	73
3D face manipulation	80
Overview	88
<i>Chapter 2: Perceptual asymmetries in judgements of facial attractiveness, age, gender, speech and expression.</i>	<i>91</i>
Introduction	92
Method	97
Subjects	97
Manufacture of stimuli	97
Description of stimuli	98
Manufacture of symmetrical face images	101
Manufacture of left-right chimaeric face images	102
Procedure	103
Results	103
Discussion	104
Expression, gender and age	105
Attractiveness	105
Lip-reading	107
Hemispheric asymmetries in processing	109

<b>Chapter 3: Facial attractiveness and symmetry.</b>	<b>112</b>
<b>Introduction</b>	<b>113</b>
<b>Introduction</b>	<b>113</b>
<b>Experiment 1 blended texture symmetry and mouth manipulation</b>	<b>120</b>
Aims	120
Methods	120
Results	123
Preliminary discussion	124
<b>Experiment 2 original texture</b>	<b>126</b>
Aims	126
Methods	126
Results	127
Preliminary discussion	128
<b>Experiment 3: The influence of self perceived attractiveness on selectivity</b>	<b>131</b>
Aims / hypothesis	131
Methods	132
Results	133
Preliminary discussion	134
<b>Experiment 4: Malaise and selectivity</b>	<b>136</b>
Aims / hypothesis	136
Methods	137
Results	138
Preliminary discussion	139
<b>General chapter discussion</b>	<b>140</b>
<b>Chapter 4: Manipulation of visual cues to perceived health in faces.</b>	<b>143</b>
<b>Introduction</b>	<b>144</b>
<b>Experiment 1</b>	<b>148</b>
Methods	148
Results	150
Discussion	150
<b>Experiment 2</b>	<b>151</b>
Method	151
Results	155
<b>Discussion</b>	<b>159</b>

<b>Chapter 5: Facial reactions to tastes.</b>	<b>162</b>
<b>Introduction</b>	<b>163</b>
<b>Methods</b>	<b>165</b>
Video-capture	165
Taste stimuli	166
Procedure	166
Video-editing and presentation	167
Video-rating	168
<b>Results</b>	<b>168</b>
Hedonic value	168
Taste strength	169
Taste judgements	169
<b>Discussion</b>	<b>170</b>
<b>Chapter 6: A test of perception of neutral expressions: normal and depressed subject performance.</b>	<b>175</b>
<b>Introduction</b>	<b>176</b>
<b>Methods</b>	<b>182</b>
Experimental set-up	182
Experimental testing	185
<b>Results</b>	<b>188</b>
Overall differences from neutral	188
Differences between High and Low BDI subjects	188
Consistency	189
<b>Discussion</b>	<b>190</b>
<b>Conclusions</b>	<b>193</b>
<b>Appendix</b>	<b>195</b>
<b>Chapter 7: Effects of adaptation to anti-face expressions on perception.</b>	<b>196</b>
<b>Introduction</b>	<b>197</b>
<b>Methods</b>	<b>200</b>
Subjects	200
Materials	200
Practice	201
Part 1: Initial testing and robustness to change in face identity	202
Part 2: Masking and delay	202
Part 3: Recognition of anti-expressions and debriefing	203
<b>Results</b>	<b>203</b>
Part 1: Initial testing and robustness to change in face identity	203
Part 2: Effects of masking and delay.	204
Part 3: Debriefing	207

<b>Discussion</b>	<b>207</b>
<b><i>Chapter 8: Thesis discussion.</i></b>	<b>212</b>
<b>Summary</b>	<b>218</b>
<b><i>References</i></b>	<b>219</b>
<b><i>Appendix to Chapter 2: Colour caricaturing of facial attractiveness.</i></b>	<b>239</b>
<b>Background</b>	<b>240</b>
<b>Methods</b>	<b>241</b>
Stimulus manufacture	241
Stimulus presentation	242
Subjects	242
<b>Results</b>	<b>243</b>
<b>Discussion</b>	<b>244</b>
<b><i>Appendix to Chapter 3.</i></b>	<b>246</b>
<b><i>Appendix: Facial affect perception in alcoholics.</i></b>	<b>251</b>
<b>Introduction</b>	<b>252</b>
<b>Method</b>	<b>255</b>
Participants	255
Materials	256
Procedure	258
<b>Results and discussion</b>	<b>259</b>
BDI score	259
Accuracy and perceptual-sensitivity	259
Perceptual-sensitivity	262
The control test	263
<b>Discussion</b>	<b>263</b>
<b><i>References for appendixes</i></b>	<b>268</b>



## **Abstract**

The face is of unparalleled importance in communication, containing cues used not only in identity recognition but, also for the assessment of character, mood, health and attractiveness. Computer graphic image (CGI) manipulation has enabled the effects of facial cues on perception to be studied from cognitive neuroscience and evolutionary psychology perspectives. A set of studies employing novel computer graphic methods to investigate facial expression, symmetry and dynamic cues related to taste is presented in six experimental chapters (2-7).

In Chapter 2 novel photo-realistic stimuli are employed to study the perceptual lateralization of facial cues for perceptions of age, gender, attractiveness, expression and lip-reading. Results suggest a right hemisphere lateralization for all perceptions except lip-reading, which appears left lateralized.

Previous studies with photographic and CGI manipulations have implied that humans unlike other animals prefer asymmetry in attractiveness judgements. In Chapter 3, new, more appropriate, CGI techniques were applied to investigate facial symmetry preference. In a series of experiments humans were found to judge more symmetrical faces as more attractive and possible individual differences in symmetry preference strength were investigated.

CGI techniques have enabled consistent qualities related to attractiveness and age to be captured from groups of face images and subsequently manipulated. In Chapter 4, these techniques are applied to capture and manipulate qualities associated with perceived skin health.

Chapter 5 represents a foray into dynamic cues related to food consumption using video. Possible facial cues to the strength, taste and the hedonic value of flavours that an observed individual was consuming were investigated.

Chapter 6 presents a novel test investigating individual differences in the percept of neutral expression. To illustrate the test: when asked to make faces expressively neutral, depressed individuals chose higher levels anger and disgust compared to controls. The test used novel 'anti-face' expression stimuli. These were later used in Chapter 7 to investigate a recent finding that adaptation to the anti-faces of individuals (faces with the opposite characteristics to a particular individual), facilitated recognition of subsequently presented corresponding individuals. The presence of analogous effects for emotional expressions was found. This effect appears to be robust to changes in individual identity, pattern masking and delays of up to a second between the adaptation and test stimuli.

Overall the thesis demonstrates the use of CGI manipulation in testing hypotheses from a variety of areas within face perception and presents a number of novel techniques that may be useful in future face perception research.

## **Note on thesis publications**

Please note that Chapter 2 derives from a publication by Burt and Perrett (1997). In Chapter 3, Experiment 2 has been published in a paper by Perrett, Burt, Penton-Voak, Lee, Rowland, and Edwards (1999) and Experiment 3 involves a novel analysis of data that has been used by Little, Burt, Penton-Voak and Perrett (2001). Chapter 5 is accepted pending revision to *Perceptual and Motor Skills*<sup>a</sup>. Chapters 4<sup>b</sup>, 6<sup>c</sup> and 7<sup>d</sup> are in submission to *Perception*, *Cognition & Emotion* and *Cognition* (brief report) respectively. The appendix titled "Facial affect perception in alcoholics" (Frigerio, Burt, Montagne, Murray & Perrett) is in review in *Psychiatry Research*. These references are listed below.

- 
- <sup>a</sup> Burt, D. M., Payne, K. R., & Perrett, D. I. (accepted pending changes). Perceptual judgements of others' tasting experiences: are they enjoying their food? *Perceptual and Motor Skills*.
- <sup>b</sup> Burt, D. M., Perrett, D. I., Murahata, R. I., & Hawkins S. S. (in review). Manipulation of visual cues to perceived health in faces. *Perception*.
- <sup>c</sup> Burt, D. M., Frigerio, E., Tiddeman, B. P., Montagne, B., & Perrett, D. I. (in submission). When does the face look neutral? A novel test of facial expression perception. *Cognition and Emotion*.
- <sup>d</sup> Burt, D. M. (in submission). Effects of adaptation to anti-face expressions on perception. *Cognition*.
- Burt, D. M., & Perrett, D. I. (1997). Perceptual asymmetries in judgements of facial attractiveness, age, gender, speech and expression. *Neuropsychologia*, 35, 685-693.
- Frigerio, E., Burt, D. M., Montagne, B., Murray, L. K., & Perrett, D. I. (in review). Facial affect perception in alcoholics. *Psychiatry Research*.
- Little, A. C., Burt, D. M., Penton-Voak, I. S., & Perrett, D. I. (2001). Self-perceived attractiveness influences human female preferences for sexual dimorphism and symmetry in male faces. *Proceedings of the Royal Society of London Series B-Biological Sciences*, 268, 39-44.
- Perrett, D. I., Burt, D. M., Penton-Voak, I. S., Lee, K. J., Rowland, D. A., & Edwards, R. (1999). Symmetry and human facial attractiveness. *Evolution and Human Behavior*, 20, 295-307.

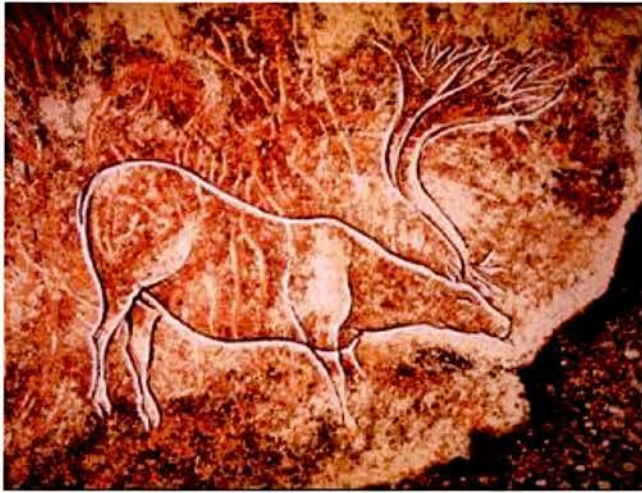
**Chapter 1: An introduction to computer graphics techniques and their uses in the study of face perception.**

## ***Introduction***

This chapter presents an introduction to the computer graphics techniques that are used in later chapters to investigate face perception. This introduction includes brief overviews of the methods used for image and shape blending, caricaturing, morphing, transforms, texture manipulation, Principal Component Analysis (PCA) of faces and 3D techniques. Following each method there are illustrations of the uses of the technique for the investigation of aspects of face perception including facial attractiveness, identity, theories of face space and the perception of facial expressions. The rest of the thesis consists of 6 experimental chapters. Each of these chapters starts with a short review of the relevant psychological literature.

The application of computer graphics to the psychology of face perception encompasses a diverse range of topics. This chapter attempts to give a selected coverage of this diversity. Attempts have been made to keep work related to the same area of psychological study together whilst, at the same time, keeping to the chronology in which the different computer graphic methods evolved. This has been difficult since the different areas of face perception that are reviewed use different methods but have been studied in parallel.

The first section of this chapter starts by looking at caricaturing since one of the technical bases of almost all of the work presented here is a framework of 'feature points' that are used to define the shape of the face and its features. The framework's first psychological use was as a tool in the investigation of the mental representation of faces for identity recognition. Caricaturing, a technique that exaggerates salient traits, was first developed at a computation level using this framework of feature points.



**Figure 1.1 Drinking reindeer from Les Combarelles, France painted C12,000 -10,500 BC (Gallery of Prehistoric Painting, New York City).**

## ***Line-drawn computer caricatures***

From early pre-history, artistic depiction has been dependent on what the artist remembered and chose to display.

For example, in Figure 1.1, a depiction of a reindeer drinking, estimated to have been painted

between 12,000 and 10,500 BC, the antlers can be seen to be larger than they probably were in reality, at least for similar species today. This might be because the antlers are a particularly salient, memorable and interesting part of the reindeer for the artist and so, possibly unintentionally, became exaggerated in the artist's work. Other prehistoric artists almost certainly tried to emphasise features of significance in their

portrayals, for instance in fertility symbols such as the Aphrodite of Willendorf, dated between 28,000 and 25,000 BC depicted in Figure 1.2.



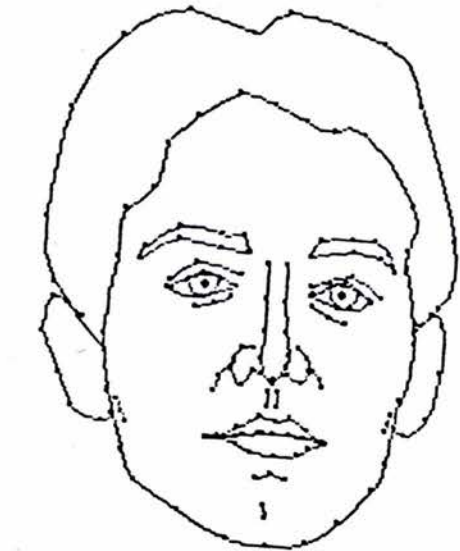
**Figure 1.2 The Aphrodite of Willendorf dated between 28,000 and 25,000 BC (Natural history Museum, Vienna).**



In more recent times the exaggeration of traits peculiar to an individual has been used in comedy, for example, in the puppets of the 1980s British comedy series *Spitting Image* (Figure 1.3). Nevertheless, the importance of distinctive traits goes beyond art and comedy. Distinctive faces have been observed to be recognised better than less distinctive faces (see Rhodes, Brennan, & Carey, 1987 for review). This led to the proposition of the distinctiveness hypothesis, namely, that faces are recognised by their distinctive features and, to the caricature hypothesis, that the distinctive traits



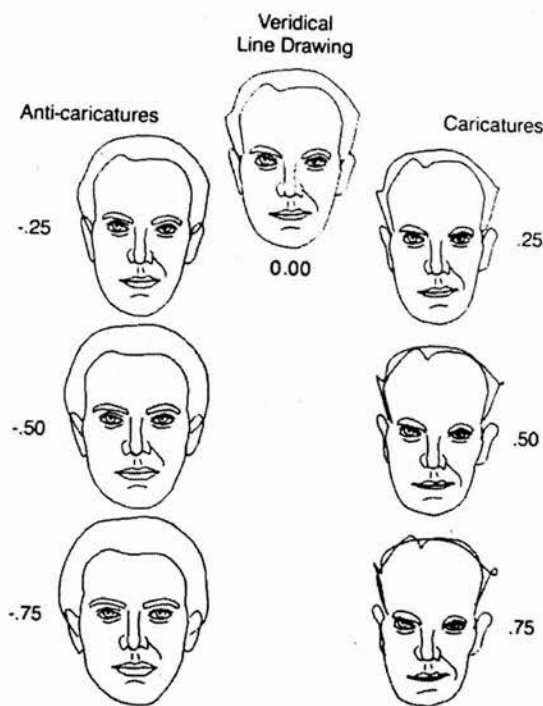
**Figure 1.3** A modern caricature the Right Honourable Margaret Thatcher as seen by *Spitting Image* (© *Spitting-Image Productions*).



**Figure 1.4** The set of points used by Brennan joined up with lines to make a line drawn sketch of the face (from Rhodes et al. 1987).

about faces may be stored in memory in an exaggerated form (Hagen & Perkins, 1983). The caricature hypothesis suggests that caricatures of individuals may be recognised better than original faces. These hypotheses were investigated using caricatures generated by Brennan's Caricature Generator (Brennan, 1982; 1985).

The 'Caricature Generator' makes use of a set of 169 feature points that are joined together to form a simple line



**Figure 1.5** The non-average attributes related to the shape of a face (example face, top) can be accentuated by caricaturing (right) to make the face shape less average, or anti-caricatured (left) to make the face shape more average. Numbers refer to distance from the veridical, with the distance from the average to the veridical being 1. (From Rhodes et al. 1987)

drawing of a face (Figure 1.4). The positions of these feature points (e.g. the left corner of the lips) on the image are recorded. The feature positions gathered from a number of face images can be averaged together to make an average face shape (see the 2<sup>nd</sup> row of Figure 1.14). Differences between any of the shapes derived from individual photographs and the average shape can be enhanced to create a caricature or diminished to create an anti-caricature (Figure 1.5).

Caricaturing and anti-caricaturing is performed by applying the function:

$$Newshape = (Veridical - Average) * Level + Veridical$$

In the above function: Newshape, Veridical and Average are all sets of points representing the new shape, veridical<sup>1</sup> shape, and the average shape of a group of faces respectively and Level relates to the level of caricature (or, if negative, anti-caricature). Thus by systematic, rather than artistic means, caricatures can be made.

Rhodes, Brennan and Carey (1987) presented student subjects with line drawings of their fellow students and university staff made using the caricature generator. The drawings were presented in their veridical form, and in both

<sup>1</sup> Veridical refers to the original, unchanged shape (or colour) derived from the face image.

caricatured and anti-caricatured forms (Figure 1.5). The level of caricaturing did not have a significant effect on subjects' accuracy in recognising the individuals but it did affect both reaction time and how good the drawings were judged to be as portrayals of the individuals. From the ratings of the line drawings it was deduced that application of a small amount of caricature (by interpolation of the data, approximately 16%) would be the optimal level of caricature for the realistic portrayal of the individuals depicted. Although the subjects were not more accurate at judging the identity of the faces portrayed in the caricatures, they were quicker when presented with caricatures at identifying the individual portrayed than when presented with veridical drawing. (Anti-caricatures were identified slower than the veridical face shapes.) These results were interpreted to be in line with the distinctiveness hypothesis and as evidence of the caricature hypothesis: caricatures are more quickly interpreted than the veridical face shapes because these caricatured shapes are more akin to what is stored in memory.

The caricature generator has also been used to find a caricature advantage (quicker or better recognition of caricatures) for other classes of objects that share analogous sets of features. It was found that caricaturing has the effect of speeding reaction time when an object is caricatured against an appropriate norm, such as a particular bird caricatured against the bird average (Rhodes & McLean, 1990). This finding was interpreted as an indication of there being different norms for different object classes.

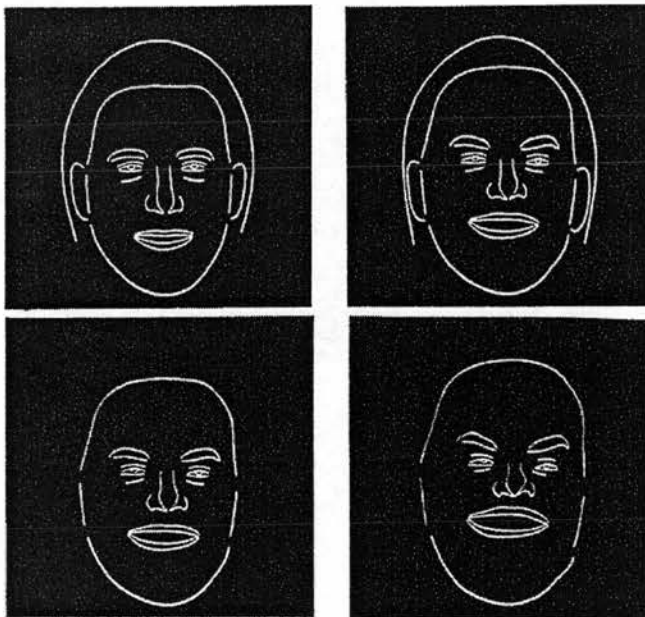
The technique of image warping (see Image Warping section below) has also been used to 'clothe' the caricatured face shapes and make the face shapes appear more realistic. Computer graphics were used (Benson, 1992; Benson & Perrett, 1991b) to warp the original images of 7 famous individuals to -32%, -16%, 0%,



+16%, and +32% shape caricature levels. To compensate for possible artefacts produced by the rendering process, the veridical (0%) image was manufactured by transforming the image first to +16% and then back again to the veridical shape (0%). In one experiment, subjects were found to rate the veridical face stimulus as being the best representation of the individual in comparison to other levels of caricature presented to them. A second experiment involved a recognition task. During each trial subjects were presented with one the 7 famous individuals' names followed by the stimulus face. An additional level of caricature, +42% was included, so 6 levels of caricature were presented on 7 individuals' face images with 7 different combinations of name label. The subject then pressed a 'yes' or a 'no' key depending upon whether they thought that the name matched the face. The +16% and +32% levels of caricature were both found to produce significantly faster reaction times and subjects were found to be most accurate at performing the task when +16% caricatures were presented. Interestingly, during this experiment a caricature advantage in reaction time was only found for trials in which the stimulus face did not match the name presented. This may be interpreted as evidence against a caricature hypothesis, since the caricature hypothesis would suggest that caricatures should be recognised more quickly, because caricatures should fit the internally held description of individual faces better. Instead this evidence supports a theory in which the caricature advantage is related to exaggerated characteristics making faces more easily dismissed as not matching because the non-matching traits are more salient. During the experiment the image of each of the 7 individuals was presented a large number of times. It is possible that the shortening of reaction time produced by caricaturing is related to subjects learning to recognise qualities of the caricatured images more quickly since the images were seen

a large number of times during testing, rather than the effect being due to spontaneous recognition of the face pictured in the image.

Benson and Perrett (Benson, 1992; Benson & Perrett, 1994) presented 3 studies in which the caricature effect was investigated for the line drawings of the faces of famous individuals. In the first study an interactive method was used to enable subjects to choose the level of caricature that represented the best likeness of the individual. During each trial subjects were presented with a line drawn face together with the name of the famous individual on screen. Mouse-movement of a 'slider' altered the level of caricature present in the line drawing by  $\pm 150\%$ . The position of the centre of the scale was varied between 0 and 100% caricature



**Figure 1.6 Examples of face shapes that were presented to subjects by Benson and Perrett (1994) with external features (top row) and without (bottom row) in veridical (left) and caricatured (right) configurations (From Benson and Perrett 1994).**

depending upon the trial<sup>2</sup>. To make the face shapes look most like the famous individuals, subjects, on average, chose to caricature the faces by +42%, significantly more than for the veridical view, confirming the results of Rhodes *et al.* (1987). In a third experiment the authors tested whether recognition was

<sup>2</sup> The bias in the centring of the scale makes the interpretation of the caricature effect more difficult since random choice along the slider would lead, on average, to a +50% caricature being chosen. The authors note that in their experiments using this method the position of the centre of the scale has non-significant effect in the ANOVA and thus argue the position of the centre of the scale has no effect on their results.

better for the famous line drawn faces (with external features present, see Figure 1.6 top row) at the average level of caricature chosen by subjects as the best likeness. Benson and Perrett found both a shortening of reaction time (veridical stimuli 6000ms; caricature stimuli 3830ms) and a gain in accuracy (veridical stimuli 50%; caricature stimuli 64% correct) in naming the caricatures of famous individuals. The authors also investigated the caricature effect for line drawings involving only the internal facial features. Previous studies have examined the caricature effect for line-drawn faces in which external features (ears and hair lines) were present. These features are not strictly part of the face, so an experiment was performed to investigate the effect of removing these external features on the level of caricature selected by subjects as most representative of the individual. Subjects were presented with stimuli without external feature lines (see Figure 1.6, bottom row), using the same methodology as employed in the first experiment to enable subjects to choose the best likeness. It was found that subjects selected a 0% caricature as being most representative of the original famous faces. The authors conclude that the lack of a caricature preference is due to the absence of a framing effect from the external facial features. This may seem an unlikely reason, especially since the face is presented with the chin and hairline as a boundary line that may be judged to frame the face. It is notable that the authors found similar effects when they presented photographic images of faces that had been shape caricatured. These images were also viewed with the external features masked out. The shape caricatured images were not chosen as the best likeness of the famous individuals pictured, but were recognised faster by subjects. The findings may instead suggest that there are different processes underlying the different optimal levels of caricature for best likeness and for fastest



**Figure 1.7** An example of a colour caricature. The RGB colouration of pixels relating to corresponding areas of the face image has been caricatured by 40% compared to average (from Lee, 1997).

recognition. The processes that mediate the best likeness effect may tend to use the external features if present.

Caricaturing facial colouration (Figure 1.7) has also been investigated and found to produce higher recognition accuracy, thus demonstrating a role for colour in the recognition of different individuals' identity in face images (Lee, 1997; Lee & Perrett, 1997). Lee, Byatt and Rhodes (2000) have recently used shape caricaturing and warping techniques to investigate models of face space<sup>3</sup>. Work on theories of face space had suggested that faces lying in less populated areas of face space may be more distinctive and recognised faster. Studies, including those reviewed above, have found that caricatures can be recognised faster than veridical faces. To test whether this effect might be due to caricatures lying in less populated areas of face space, caricatures, veridical and anti-caricatured faces of well-known individuals were first rated by a set of subjects to provide distinctiveness ratings. Multidimensional scaling techniques were then applied to the data to produce an optimally fitting model of

---

<sup>3</sup> Face space refers to a theoretical, multi-dimensional space in which faces are distributed. The axis of the multiple dimensions relate to the ways in which the faces differ from one another and the spacing of the faces along the axis relate to how much the individual faces differ from one another along each axis.

psychological face space. Caricatures were rated as more distinctive and were found to occupy regions of face space that were less densely populated. The model also had predictive value in predicting recognition accuracy of the faces. Notably, for work on attractiveness (see below), the image of the average face (manufactured by blending face images together) was found not to be located at the centre of the face space, described by the model, as might have been expected from an average. The authors conclude that this is possibly due to the average face's smooth texture<sup>4</sup>.

---

<sup>4</sup> This finding that the manufactured average face was not found to be at the centre of face space (where the average would be expected to be found) is of interest in the light of investigations into the hypothesis that averageness is attractive that are presented later. It may well be that the manufactured average is off centre because of its atypical texture.



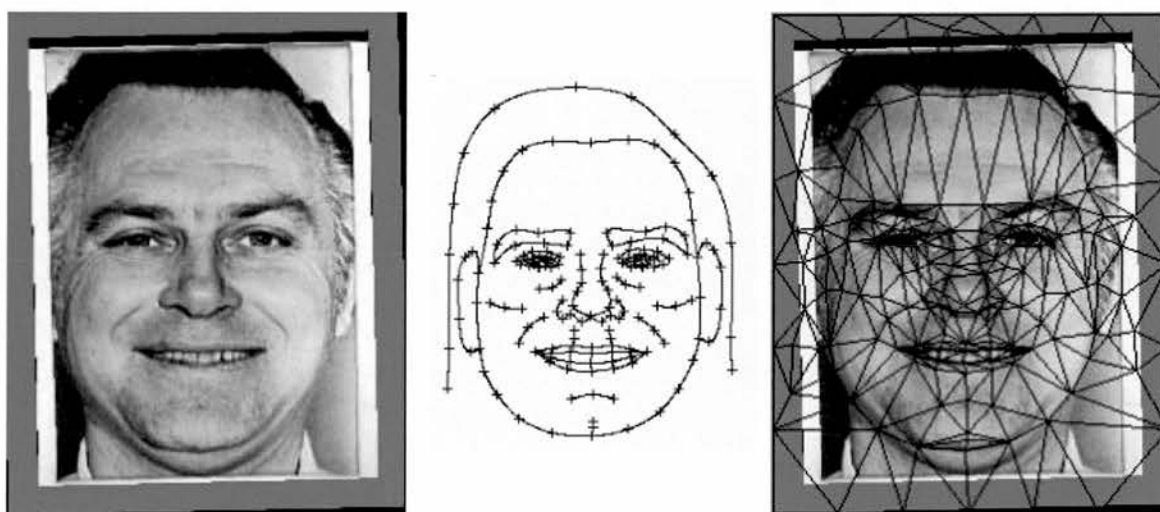
## ***Image warping***

The computer graphic methods used in most of the previously cited papers have been based on simplified line drawings made by joining up 'feature points' corresponding to reliably found features. The resulting stimuli are highly simplified, lacking many of the characteristics of real faces in real face-to-face interactions. This lack of realism may affect the results of experiments. For example, Rhodes *et al.* (1987) have suggested that the caricature effect might be a result of the lack of information contained within the line drawings. The next step towards more realistic portrayal of faces has been the use of distortion of the original image into the manufactured face shape by a process called image warping. Image warping gives rise to 'photo-realistic' stimuli possessing all of the cues that were present in the original photographic image, although not all the cues present in the image are manipulated and the cues that are manipulated may not be changed in a realistic fashion. It is also worthwhile noting that photographic images are themselves only representations of the real world scenes that they capture.

Photographic images lack realism in many ways. For instance they are not 3-dimensional and do not move. The lack of both of these features is, to an extent, addressed in this thesis and in studies performed by other groups applying computer graphics to the study face perception. However, it should be noted that the photographic image does not really interact, even when animated. The image lacks many of the other cues normal in a human-to-human interaction, the face contained in the image does not have feelings or exhibit the complex contingent behaviours that are a part of human interaction. Although, this can mean that subjects may not interact with computer graphic stimuli as they would with a real human, the use of photographic images does allow experiments to be performed with relative ease that

would otherwise be vastly expensive (plastic surgery bills etc.), immensely time consuming and rather unethical. So this first step in manipulating photo-realistic images is important.

The aim of image warping is to alter the shape of the object in original image from the configuration of the original set of feature points to a new set of feature points. Computer based images are made up of a 2 dimensional pixel matrix. In colour images each pixel has 3 numbers associated with it, the intensity of red, green and blue<sup>5</sup>. The task of image warping can be looked upon as calculating a new image (a 2 dimensional matrix of pixels) in which the colouration of each pixel is calculated from



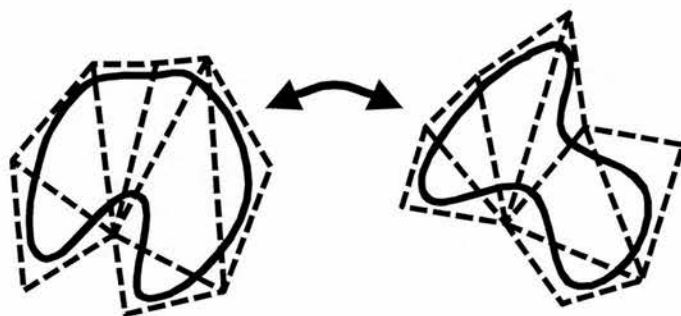
**Figure 1.8** An original image (left) is delineated to give its shape (centre). The feature points of the face shape can then be used to break up the image into a series of triangles (right). (Original images from Ekman and Friesen, 1976.)

the original image using a transformation based on the original image's feature point positions and the new feature point positions. This is normally achieved in one of several ways.

---

<sup>5</sup>Here 24bit RGB images are talked about. There are other ways of coding image colour but this is the method that is used most often. Early work with image warping was performed on grey-scale images that possess only one value, luminance, rather than 3 (red, green and blue).

The method that has been used for manipulating face images in all of the experiments performed in St Andrews is one in which the original image is broken up into a series of triangles based on the feature points (see Figure 1.8) (Benson & Perrett, 1991a; Benson & Perrett, 1991c) whose position is derived from Brennan's



**Figure 1.9 Stretching and shrinking the triangles along their vertices to distort the blob (figure credits Rowland, 1998).**

(1985) framework. Stretching or shrinking the triangles along their vertices can then be used to distort the objects in the image (see Figure 1.9). Based on one set of feature points it is possible to break the image up into triangles in many different ways. The connections between feature points can be manually defined or an automated method used. One of the best methods is based on Delaunay triangulation (Ruppert, 1995). Delaunay triangulation algorithms result in a mesh of triangles in which each triangle

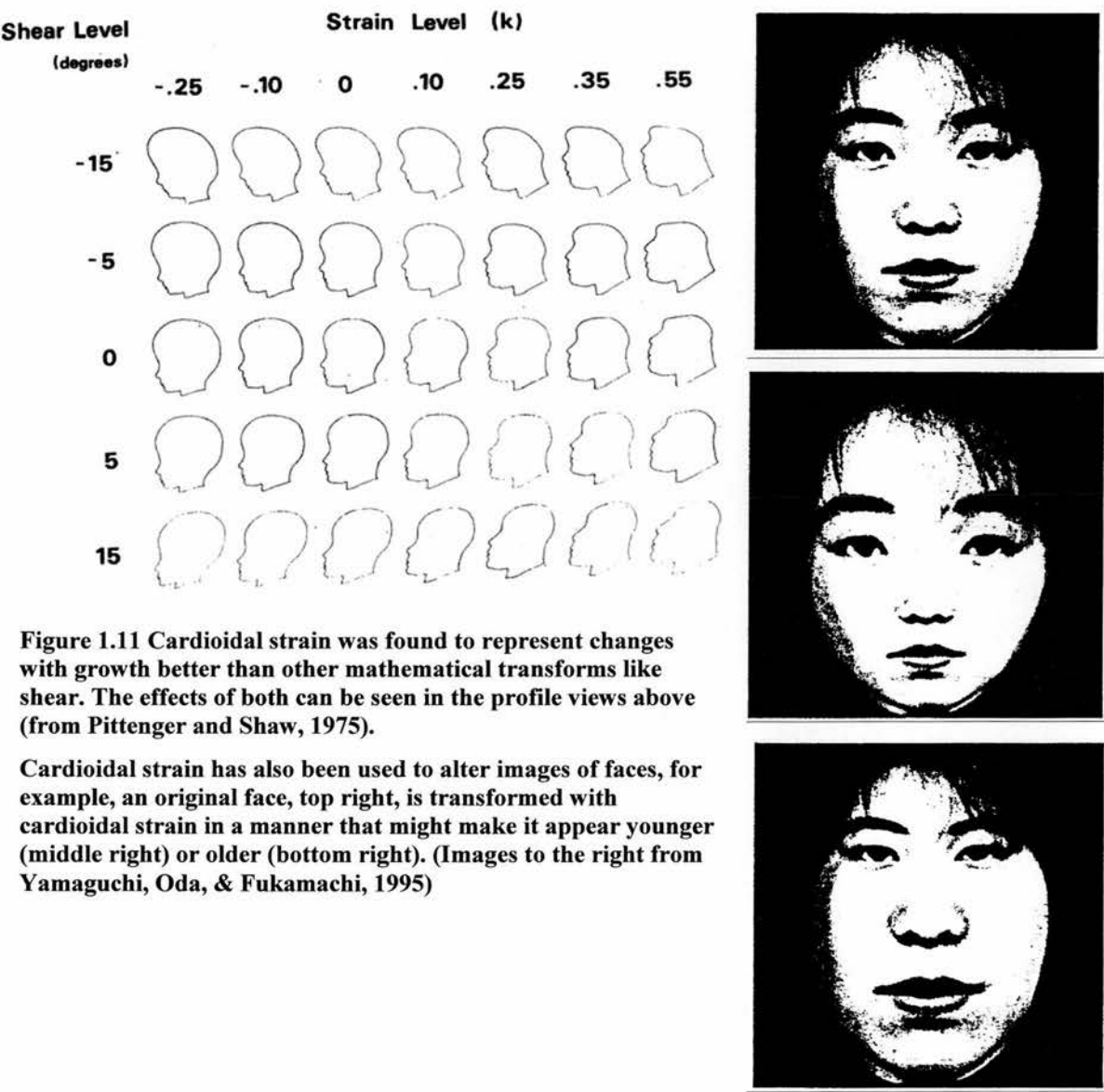


**Figure 1.10 Stretching and a triangle distorts the square (figure credits Rowland, 1998).**

is as near to being an isosceles triangle as possible. Subject to confines introduced by the configuration of the set of feature points, no points fall within the interior of any circle that passes through the 3 vertices that make up the triangle. This has the effect that none of the triangles in the mesh of triangles overlap. As with other forms of warping, this method under certain conditions produces distortions. Features that continue across triangles can become distorted when warped (see Figure 1.10) or may overlap when warped into the configuration of another face shape (see Tiddeman,



Duffy, & Rabey, 2001b for a method to avoid overlap). Other methods of image warping are based on curves drawn between the feature points (Wolberg, 1990), using fields of influence situated around feature points (Beier & Neely, 1992) and using optic flow (a method originally used to find the correspondence of features between images from an animated sequence; Beymer, Shashua, & Poggio, 1993).



**Figure 1.11** Cardioidal strain was found to represent changes with growth better than other mathematical transforms like shear. The effects of both can be seen in the profile views above (from Pittenger and Shaw, 1975).

Cardioidal strain has also been used to alter images of faces, for example, an original face, top right, is transformed with cardioidal strain in a manner that might make it appear younger (middle right) or older (bottom right). (Images to the right from Yamaguchi, Oda, & Fukamachi, 1995)

Warping has been used in conjunction with shape based caricaturing to make photo-realistic shape caricatures of face images (see previous section), and for such things as the examination of the effects of cardioidal strain, a simple mathematical model of how the shape of the face might change during growth (see Figure 1.11,

Pittenger & Shaw, 1975). Examples of a face image that has been subjected to shape changes to increase and decrease cardioid strain are presented in Figure 1.11 (Yamaguchi, Oda, & Fukamachi, 1995). Warping has also been used for the investigation of the effect of changes in the shape and size of individual features on face perception. For example, Keating (2001) studied the effect of manipulations of eye and lip size on attributions of maturity made to individuals in photographs. Warping though is most often used with other image processing techniques such as blending or transforming.

## ***Image blending using feature points***

Computerized methods of image blending derive from the earlier technique of photographic image superimposition first used by Galton (1878) to make averages from groups of face images. After projecting one portrait photograph onto photosensitive paper, Galton then exposed another over the first rotating, translating and scaling the image projected from the negative so that the eyes of the first and subsequent images match and tilting the photosensitive paper to match the face length. This process gave rise to images like that in Figure 1.12. Galton hoped that



**Figure 1.12 The photographs of 49 female undergraduates composited together by Stoddard (1886)**

his average faces would preserve the qualities of the group of individuals from which they came. In many ways they do. For example, the composite images can preserve gender. The image made of female faces by one of Galton's contemporaries (Figure 1.12) looks female whereas Galton's averages of male faces look male (Figure 1.13). Galton also thought that consistent traits to criminality would show up in an average of criminals' faces. However, his work on extracting criminal characteristics was not so successful. Galton's average criminal images (Figure 1.13) do not look very criminal but his work did lead him to observe that the average face images that he had made were "much better looking than those of the components" (Galton, 1878, p.97). More recently this effect of averaging by superimposition has been interpreted as evidence for the idea that having average qualities makes faces attractive (Langlois & Roggman, 1990; Langlois, Roggman, & Musselman, 1994; Langlois, Roggman, & Rieser-Danner, 1990).




HEALTH.	DISEASE.	CRIMINALITY.
		
23 Cases Royal Engineers 12 Officers 11 Privates	6 Cases  9 Cases Tubercular Disease	8 Cases  4 Cases 2 Of the many Criminal Types

Figure 1.13 Galton made blends of different groups of faces in an attempt to produce the average upstanding citizen, criminal type, and diseased individual's typical appearance. (Images from Galton, 1878)

There are good theoretical reasons for believing that we evolved to choose the average of the population as more attractive than individuals far from average. The average is by definition free from abnormalities, so choosing average may be a good way to avoid genetic defects and diseases. The average individual in a population is also likely to be more heterozygous (have a more diverse mixture of genes) than less typical individuals, and so may be less susceptible to problems related to inbreeding (Thornhill & Gangestad, 1993 for discussion). Individuals with a more average appearance may also be healthier. Rhodes et al. (2001b) found that a more average appearance (lower rating for facial distinctiveness) was related to the current health of 17-year-old female individuals and, for 17-year-old males, a more average appearance was related to better childhood health. Representations of averages may also be used by the visual system in recognition. For example, theories of object recognition, which include faces as a class of objects, are often based on differences from a category norm or average (e.g. Rhodes & McLean, 1990). If these representations of averageness exist, they may be applied to mate selection. Run away selection for costly exaggerated traits (Fisher, 1930) can also be avoided by a preference for average.

Langlois *et al.* used computer graphics techniques to average face images by rotating, scaling and translating the images so that their eyes were aligned prior to

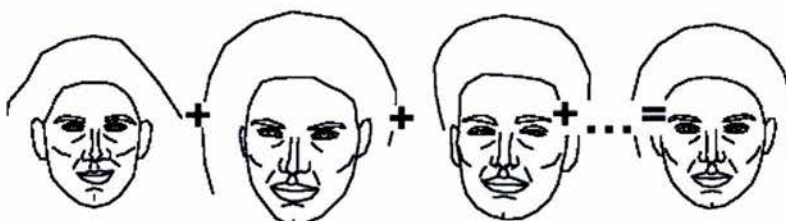
superimposition. Superimposition of greater numbers of faces was found to produce more attractive face images, leading them not only to conclude, but also to name their paper, “Attractive faces are only average” (Langlois & Roggman, 1990).

The methods that both Galton and Langlois *et al.* used to create averages, result in an image that is the mathematical average and, in that way, representative of the group of *images* from which it was formed. The face in the average image is though uncharacteristic of the faces that went into making the average image. Average face images are blurred and free of the imperfections present in the original face images from which they were created. Average images are thus not typical of the face set. The blurring that is induced by computer averaging techniques is complex and, as we will see in a later section relating to texture, applying a simple ‘sharpening filter’ does not reverse this blurring produced by averaging. To test the hypothesis of Langlois *et al.*, that the average face is most attractive in terms of shape, Perrett *et al.*



**Figure 1.14** Blending the image on the top line without alignment produces a blur.

If the face shapes are delineated an average shape can be formed (middle line)



All images can then be warped into this shape and blended together to produce a clearer average image (bottom line).







**Figure 1.15** Ways to blend Margaret Thatcher. Six original photographs (a-f) of Margaret Thatcher where blended together using three different techniques: (g) superimposing without normalising. Aligning the eyes in a way analogous to Langlois et al. (1990) (h) and using feature points to align the feature positions. (images from Benson & Perrett, 1993).

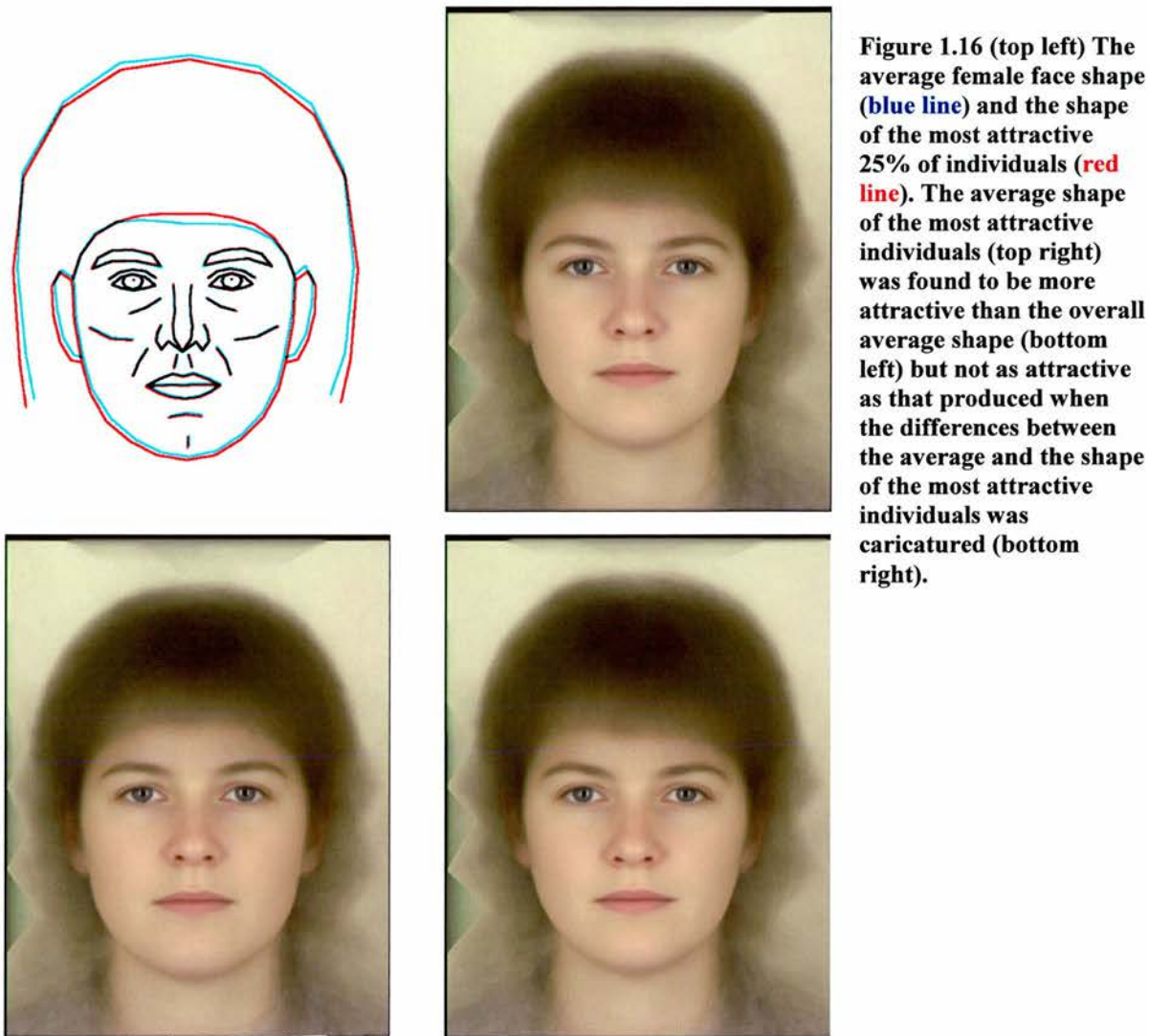
(1994) used the computer graphics techniques of shape averaging and caricaturing, in combination with image blending, to produce a set of stimulus images in which texture is controlled.

Benson and Perrett (1991a; 1993) found that if images were first warped into a common shape defined by feature points, they could then be combined by cross dissolving to form an average (Figure 1.14). This technique produces average images,

which, although slightly blurred, appear to preserve many of the qualities common to the group of images from which they were derived (Figure 1.15).

If, as suggested by Langlois and Roggman (1990), the single quality that is attractive in faces is their averageness, then the average shape of a population of faces should be as attractive as the average shape produced from a subset of the most attractive faces from the population. Perrett *et al.* (1994) tested whether subjects found the average face shape of the whole population or the face shape made from the most attractive individuals most attractive. In the experiment subjects first rated the attractiveness of 60 female faces. Using similar methods to those used by Brennan (1985), the average shape for the 15 most attractive faces (Figure 1.16 top left red lines) and all 60 of faces in the group (Figure 1.16 top left blue lines) were calculated. Both of the shapes were then ‘clothed’ with identical colouration and texture

manufactured by blending all the individual images together into the destination shapes (Figure 1.16 top right and bottom left). Rather than preferring the average



shape produced from the whole population of faces, subjects preferred the average shape made from just the most attractive faces. Furthermore, when the differences between the average of the population and the average made for the most attractive faces were caricatured and clothed with average colouration to produce a +50% caricatured high average face (Figure 1.16 bottom right), this stimulus was preferred over the other stimuli by subjects. This finding demonstrates that the differences in shape between the population and the most attractive individuals can be manipulated using photo-realistic images. Similar results were also found for male faces, although

caricaturing the high stimulus did not result in a stimulus that was significantly preferred over the high stimulus. This smaller, non-significant effect may be a reflection of a lack of agreement between subjects on cues to male facial attractiveness<sup>6</sup>.

Rubenstein, Langlois and Roggman (2001) tried to replicate the study by Perrett *et al.* (1994), but found that the average of the most attractive faces was as attractive as a average made from a cross-section faces of varying attractiveness. Their result may be due to the less sensitive techniques adopted for acquiring subject preferences or the methods of graphic manipulation used. Nevertheless, Rubenstein *et al.* suggest (amongst other things) that the results of Perrett *et al.* may be due to the more attractive faces being more average. They also define what they mean by average stating that their term ‘ “average” (more accurately referred to an “averaged” face) refers to the physical configuration of faces created by averaging multiple individual faces together mathematically.’ and that ‘a mathematically average face is not an average or a common face.’ (p.11). This insistence on a procedural definition of average may seem strange when making a theoretical argument. This definition sits poorly with the arguments above for averageness being preferred, since the arguments above are related to the individual in the image being average rather than, as here, the image itself being processed using a specific computer graphic manipulation. It is notable that the computer graphic manipulation is known to rid the face image of minor individual characteristics and abnormalities typical of the set of images from which the average is derived, possibly leading to the manufactured average face image not being in the centre of face space (Lee *et al.*,

---

<sup>6</sup> Comparable effects can also be found when colour alone is manipulated (see Chapter 1 Appendix).



2000) and appearing to be younger than the face images from which it is manufactured (e.g. Burt & Perrett, 1995)<sup>7</sup>.

Examination by eye of the stimuli used above by Perrett *et al.*, suggested that the more attractive female face looked more feminine, and possibly younger having a smaller chin and larger eyes relative to their face size (Figure 1.16); whereas it was more difficult to propose what the differences between the face shapes of the male averages were. There is substantial evidence that feminine features, such as large eyes and an expressive mouth, are attractive in female faces (such as Cunningham, 1986). There is also evidence that faces with masculine features, such as a large jaw and prominent brow ridge, are found attractive in male faces (such as Cunningham, Barbee, & Pike, 1990; Scheib, Gangestad, & Thornhill, 1999). Cunningham *et al.* (1990), suggested that females might, in their choice of male faces, be attracted to masculine features but also to feminine features. This, Cunningham *et al.* (1990) proposed, was a possible reflection of mixed motives in choosing a mate. These 'mixed motives' might relate to the lack of a significant effect on attractiveness found, for male faces, in the comparison between the average shape of the most attractive faces and its caricature (Perrett *et al.*, 1994). To understand further the impact of sexual dimorphism in face shape on attractiveness, the shape of male and female

---

<sup>7</sup> Later studies using image blending techniques, similar to those used by Perrett *et al.* (1994), have also found that average faces tend to be more attractive than their components (in adults, Little & Hancock, In press; Rhodes & Tremewan, 1996; and infants, Rubenstein, Lanlgois, & Kalakanis, 1999). In some of these studies noticeable abnormal facial features e.g. spots have been removed prior to the blending of the face images and thus have been shown not to cause the heightened attractiveness of average face images.

average images were manipulated to make them more or less masculine and subjects' responses investigated (Perrett et al., 1998; Rhodes, Hickford, & Jeffery, 2000).

Perrett *et al.* (1988) found that, when asked to select the most attractive female face shape, subjects chose a more feminine face shape than the average, but, when asked to choose the most attractive male face shape, subjects also chose a male shape face that was slightly feminised compared to the average. To understand further the reasons for this, other subjects were asked which of the male faces, the average, 50% feminised or 50% masculinised, they rated higher on a number of traits. The more masculine male face was rated as looking more masculine, dominant and older but was also rated as less of a good parent, less honest, less warm, less cooperative and lower on emotionality than the less masculine looking faces. These attributions to the more masculine male faces may be related to more masculine males' behaviour. For instance, studies have found that men with higher testosterone levels tend to have more marriage break-ups (Booth & Dabbs, 1993).

Testosterone has been suggested as an immune suppressant that is only tolerated at high levels by healthy individuals. Since high levels of testosterone produce more masculine facial features, it has been suggested that these may be an honest indication of health (see Thornhill & Gangestad, 1999 for review). The cost of high levels of testosterone would be expected to confer benefits to the organism, but it appears that females have less of a preference for the more masculine facial features, that may be a sign of high testosterone. Why then, do the genes related to high testosterone levels not die out? These genes might, in some way, have preferential transfer to the next generation. This transmission to the next generation may relate to male individuals, with more masculine features, tending to be more dominant over other males and so rise in the social hierarchy and gain better access to females.

Alternatively, females may in some way prefer more masculine male features, such as when conception risk is highest.

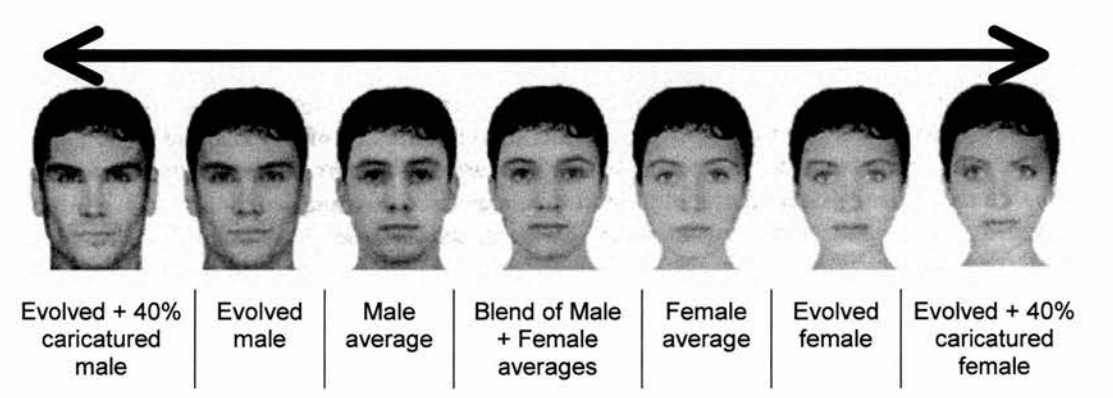
To test whether female preferences might shift in favour of more masculine faced males when conception risk is high, subjects in the UK and Japan were asked on multiple occasions through their menstrual cycle to select their most preferred male face from several continua that ranged in shape from 50% more masculine to 50% more feminine than the average male. For each subject the date of onset of their next menses was recorded enabling the dates of the high and low conception risk phases of their menstrual cycles to be ascertained. In both UK and Japanese populations, female subjects selected slightly feminised versions of the male faces as being most attractive. At times when subjects were classed as being at high conception risk they selected significantly more masculinized male faces than at other times (Penton-Voak *et al.*, 1999a). This effect was most potent when subjects were cued to assessing attractiveness in the context of a short-term relationship. A results that was interpreted as reflecting a tendency for females to prefer more masculine males during times of greater conception risk. More recently Johnston *et al.* (2001), using different methods, has replicated this relationship in which female subjects during higher conception risk periods of their menstrual cycle select more masculine faces.



**Figure 1.17** A mockup of the face prints program in action. (Image from Johnston, 1994)

In their experiment, Johnston *et al.* (2001) repeatedly presented subjects with a single continuum that ran from highly masculine to highly feminine, using a similar

presentation technique to Penton-Voak *et al.*. The continuum was manufactured using a combination of techniques: morphing, caricaturing and an ‘identikit kit’ method called FacePrints that enables subjects to evolve faces using a database of face parts (Figure 1.17, Johnston, 1994). Images from the continuum are illustrated in Figure 1.18. The female image to the right of centre of the continuum (*Female average* on Figure 1.18), was produced by blending together 16 female face photographs taken of Caucasian individuals by the artist Gomi (1988). The male average (*Male average*) was manufactured from the photographs of 16 male college students. The faces towards the extremes of the continuum from the male and female averages were constructed in a previous experiment using Johnston’s computerized identikit kit. In this previous experiment, subjects had been asked to evolve either a masculine or a



**Figure 1.18 Selection of frames from the facial morph movie. Running from left to right: caricature of the evolved masculine face by 40%, masculine face evolved using FacePrints, average male student, male and female average, average female from Gomi (1988) images, evolved feminine face, caricatured evolved feminine face by 40% (from Johnston et al. 2001).**

feminine face. These masculine and feminine identikit faces were then shape caricatured by 40% against the average (formed by blending of the *Male* and *Female* average) to produce the extremes. Morphing was then used to interpolate frames between these pairs of images.

In the same paper, Johnston *et al.* (2001) also report that the most attractive male face was a slightly more masculinized than the average male. The methods used by Johnston *et al.* are different from those used in other studies (Penton-Voak *et al.*, 1999a; Perrett *et al.*, 1998; Rhodes *et al.*, 2000). In these other studies a preference was found for a more feminine male face shape than the average male. Johnston *et al.* criticise the previous studies as flawed as they are based on averaging and caricaturing techniques, but employs both techniques, adding a third to make the single continuum that is use for the above experiment. The preference that they identified for masculinity may be, as they claim, just that, or may be an artefact of their methodology.

Artefacts may result because the more masculine face image that Johnston *et al.* use is not from the same set as their average male student face and so may not be comparable. Furthermore, subjects may be conscious of the dimension that is being manipulated since they are given multiple trials with a variety of instructions that may cue subjects to the nature of the continuum (find the average face, the average male face, the most attractive male face, the most masculine male face, ... etc.). An alternative explanation of the difference between their results and previous studies may be that rather than manipulating shape alone, Johnston *et al.* morph the images, altering both colour and shape cues at the same time. This has the effect that trying to select a very feminine male face will result in a female face being chosen rather than a female shape, with male colouration, as in other studies (Penton-Voak *et al.*, 1999a; Perrett *et al.*, 1998; Rhodes *et al.*, 2000). The experimental result, that a masculine male face is preferred could, be explained by avoidance of the female faces in the continuum.

Individual differences among subjects in what is found attractive have also been investigated using caricaturing. Little *et al.* (2001)<sup>8</sup> investigated the differences in perceptions of female subjects depending upon self-ratings of their own attractiveness. Female subjects who rated themselves as highly attractive selected more masculine male face shapes (than did female subjects who rated themselves as average or low) as being most attractive from a continuum running from masculine to feminine male faces<sup>9</sup> when asked in the context of long-term, but not in the context of a short-term relationship. More recently Penton-Voak, Little, Jones, Burt, Tiddeman, and Perrett (in submission) have investigated the effects of a physical measure related to attractiveness, waist to hip ratio (WHR), on preferred degree of facial masculinity. WHR has been found to relate to female attractiveness, fertility and health (e.g. Singh 1993; 1995)<sup>10</sup>. Penton-Voak *et al.* found no effect of relationship term for low (more attractive) WHR women but women with high (less attractive) WHR selected more 'feminine' looking male faces when cued to a long-term relationship than when cued to a short-term relationship. These results are interpreted as the less attractive female subjects using different mate finding strategies at different times, whereas the more attractive subjects may use one strategy all of the time. The study also found an analogous relationship between the female subjects' own attractiveness (as rated from photographs by independent observers), relationship term and selection of more masculine face shapes.

---

<sup>8</sup> This study also looks at the effect of own rated attractiveness on symmetry see Chapter 4.

<sup>9</sup> The stimuli used were the same as those used by Penton-Voak *et al.* (1999a).

<sup>10</sup> The specifics of these relationships are disputed. Tovee regards the relationship between WHR and attractiveness as being primarily mediated by Body Mass Index (e.g. Tovee & Cornelissen, 1999).



The technique of blending has also been used for the investigation of the mere exposure effect. The mere exposure effect is the tendency to preferentially like items for no other reason than having been exposed to them before (see Bornstein, 1989 for review). Rhodes, Halberstadt, and Brajkovich (2001a) tested whether the mere exposure effect would generalise from individual face images to blends composed of the same face images. During an initial exposure phase subjects were repeatedly (4 times) presented with the faces of different individuals. In the test phase subjects were then presented with averages made from groups of 6 face images, that had either previously been seen in the test phase or not, for rating on attractiveness and likeability scales. The study found that seeing the component faces of a blend in the initial exposure phase of the experiment led later, during the test phase, to increased liking and attractiveness<sup>11</sup> ratings. The study showed that the mere exposure effect generalized from individual faces that had been seen to blends of those faces. This finding is discussed in terms of the mere exposure effect being one of the factors influential in average faces being attractive<sup>12</sup>.

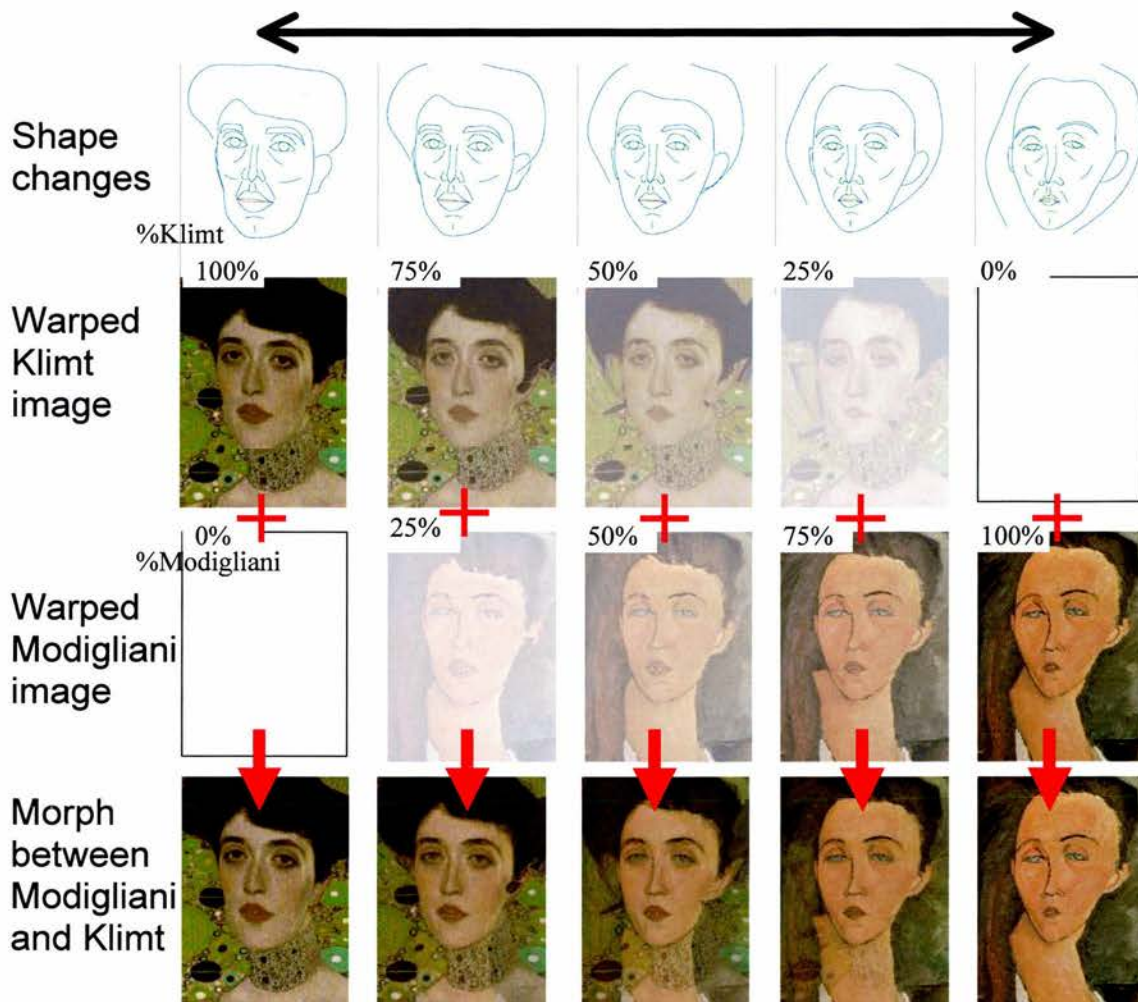
---

<sup>11</sup> The mere exposure effect for attractiveness was not replicated by a second experiment in the same study.

<sup>12</sup> Another factor that has been suggested to be influential in findings that average faces are most attractive is symmetry. The relationship between symmetry and attractiveness is investigated Chapter 3.

## Image morphing

Another computer graphic technique that has been used extensively to study face perception is image morphing. Since the start of the 1990s, image morphing has been used extensively in film and other forms of popular media both for spectacular effects and, as the required technology has become cheaper and easier to use, for



**Figure 1.19** Morphing a portrait by Klimt to a portrait by Modigliani. Both original images were first delineated. The face shape painted by Klimt is shown on the top left with his work directly under it. The shape of Modigliani's work is shown on the top right. The shape has been 'normalized' (rotated, translated and scaled so that the eye positions match Klimt's work). Modigliani's work has then been warped to these new feature positions and can be seen on the right of the third row. In the top row the position of the feature points are morphed in steps from the shape of Klimt's portrait to that of Modigliani's. In the 2<sup>nd</sup> and 3<sup>rd</sup> rows Klimt's and Modigliani's paintings are warped into the corresponding shape and faded. The images in the 2<sup>nd</sup> and 3<sup>rd</sup> rows are then added together to produce the final morph.

continuity between different sequences of film. The popularity of image morphing has meant that cheap (and even free) software to perform image morphing is now readily available. Using the modern graphics hardware typical of today's PCs, image morphing can be performed in real time (e.g. using the graphics software as displayed at the Glasgow Science Centre or Camera Obscura in Edinburgh, Tiddeman, 2001). The technique of image morphing is a combination of image warping and fading to enable a smooth transition from one image to another. The technical literature for image morphing is the same as that for image warping covered earlier, although, unlike image warping, the technique of cross dissolving between the images is also used. The process of morphing is described in Figure 1.19.

Image morphing has been used to blend pairs of images together to produce average images (as covered in the previous section), to produce symmetrical face images (Rhodes, 1986; Swaddle & Cuthill, 1995) and to alter the degree of facial expression present within an image. This section concentrates on the literature regarding morphing in two experimental areas: the investigation of categorical perception and the investigation of the perception of facial expressions.

Categorical perception refers to the finding that sensitivity to the same sized physical differences between stimuli is greater if the difference occurs at the borderline between categories, the 'categorical boundary', than if the same sized difference occurs within a category. Categorical differences have been found in many perceptual domains. For example, a pair of colour swatches that are both categorised as being green in colour may be harder to discriminate between than a pair consisting of a greenish yellow and a yellowish green swatch even if the physical, i.e. wavelength, difference in colouration of the individual swatches is the same (for

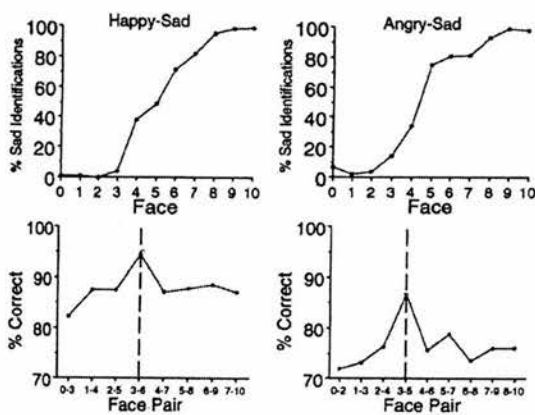


Figure 1.20 Above are examples of the computer line drawn morphs used by Etcoff and Magee (1992). Subjects completed two parts of a test. In one, the discrimination task, subjects were sequentially presented with 3 line-drawn faces. The third face was always the same as either the first or second face and subjects had to say which. The proportion of times that subjects were correct is plotted in the graphs (lower left). Later, to locate the categorical boundary between the pairs of expressions that each continuum was manufactured from, subjects were asked to choose which expression each individual line drawn face was showing choosing between the two relevant expression labels. (From Etcoff and Magee, 1992)

review see Harnad, 1987). Analogous effects have been found for the perception of different facial categories.

Etcoff and Magee (1992) used morphs based on line drawings (as used by Brennan, 1985) of the six basic facial expressions from photos by Ekman and Friesen (1976). Line-drawn morph continua, as illustrated in Figure 1.20, were made ranging between: angry and sad, angry and afraid, angry and disgusted, happy and sad, happy and neutral, sad and neutral, happy and surprised, and surprised and afraid. As can be seen from Figure 1.20, Etcoff and Magee found discrimination improved at the categorical boundaries between expressions.

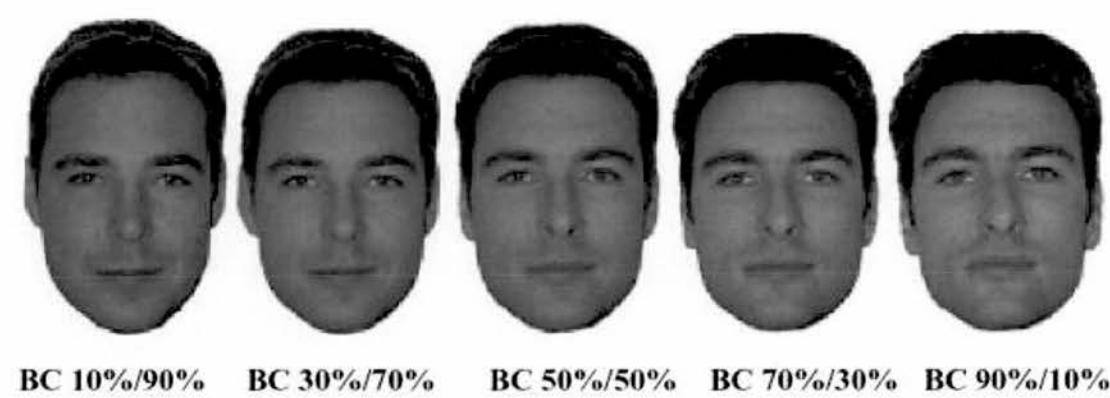
Calder *et al.* (1996a; and Young *et al.*, 1997) have extended the study of Etcoff and Magee by showing that with images of shape-morphed facial expressions, subjects perform stimulus discriminations better across the categorical boundary between expressions. The process that leads to categorical perception of facially presented emotions may develop very early in life or may possibly be innate. Kotsoni,

de Haan and Johnson (2001), investigated emotional perception in 7 month-old infants. Infants preferentially look towards fearful rather than happy faces (Nelson, Morse, & Leavitt, 1979). Using this preference, Kotsoni *et al.*, found the position of the categorical boundary between happy and fear faces by presenting pairs of morphed images from sequences running from a full blown happy expression to a full blown fear expression and analysing the infants looking time. After finding the categorical boundary, a second experiment was performed. In this experiment equidistant pairs of images from the morph sequence were displayed. These image pairs either spanned the categorical boundary or did not. The infants were found to discriminate more, in terms of their looking times, between pairs of faces that spanned the categorical boundary than between pairs of faces that did not. Thus, even young infants seem to have categorical boundaries for at least one pair of facial expressions. Intriguingly, these categorical boundaries between facial expressions may fade in old age (Bruyer & Granato, 1999).

Other researchers have used image morphing to investigate categorical perception of other facial dimensions. Beale and Keil (1995) presented subjects with famous faces that varied in the amount of different individual's identity present by using morphs that ran from one familiar face's identity to another's (e.g. President Kennedy to President Clinton). They found categorical perception for the discrimination of identity. These categorical effects for familiar face identity have also been investigated using electro-encephalogram (EEG) techniques. Campanella *et al.* (2000), familiarized subjects with multiple views of novel individuals. Subjects were then trained to name the individuals presented and the categorical boundaries for image morphs between different individual identities found using a similar method to Etcoff and Magee. While EEG was recorded, subjects were presented with pairs of



stimuli, one face at a time on screen. During this presentation subjects were asked whether each pair of stimuli represented the same individual or not. The stimuli were pairs of frames from morphed sequences between from one individual's face image and another's (Figure 1.21). In some cases the pairs of images were from different sides of the categorical boundary between the identities of two individuals. In other cases the image pairs were from the same side of the perceptual boundary. Stimuli were also included in which the same individual's face was morphed between two photographs taken of the same individual on different occasions.



**Figure 1.21** Example of morphed face stimuli going from 90% of an individual's identity to 90% of another individual's identity as used by Campanella et al. (2000). (Figure from Campanella et al. 2000)

Campanella *et al.* found that when a ‘within category’ (i.e. same individual) pair of images was presented there was an EEG response (N170) originating in the right occipito-temporal cortex, after presentation of the second stimulus. This EEG response was less evident for ‘between category’ (different individual) pairs.

Campanella *et al.* suggest that this response is reminiscent of a priming effect.

Computer graphics techniques including morphing have been used to generate faces conveying controlled amounts of emotional expression for the study of expression perception. Considerable work has been performed since 1994 on the



amygdala and its function in emotion perception using computer graphics techniques. Adolphs *et al.* (1994; 1995) suggested that two individuals with bilateral amygdala damage had particular difficulty recognising expressions of fear. These studies were later supported by work of Calder *et al.* (1996b) in which two individuals with bilateral amygdala damage were presented with faces from image morphs between different facial expressions to test emotion perception. The testing with morphed expressions was performed using pairs of faces morphs between the expressions of: happiness and surprise; surprise and fear; fear and sadness; sadness and disgust; disgust and anger; and anger and happiness. These combinations of expressions were used since they represent transitions between the most commonly confused expressions. The set of emotions and morphs presented in this test is referred to as the 'Emotional Hexagon' test since the 6 different emotions can be visualized as the corners of a hexagon, connected round its perimeter by the morph sequences.

Using the Emotional Hexagon, Calder *et al.* found that the two individuals with bilateral amygdala damage, DR and SE, both showed poorer recognition of expressions than controls. DR showed significantly poorer recognition of the expressions consistently recognised by controls as fear, anger and disgust while SE's data showed a trend towards poorer recognition of fear and anger with normal recognition of the other emotional expressions (a sample of the data from DR is presented in Figure 1.22).

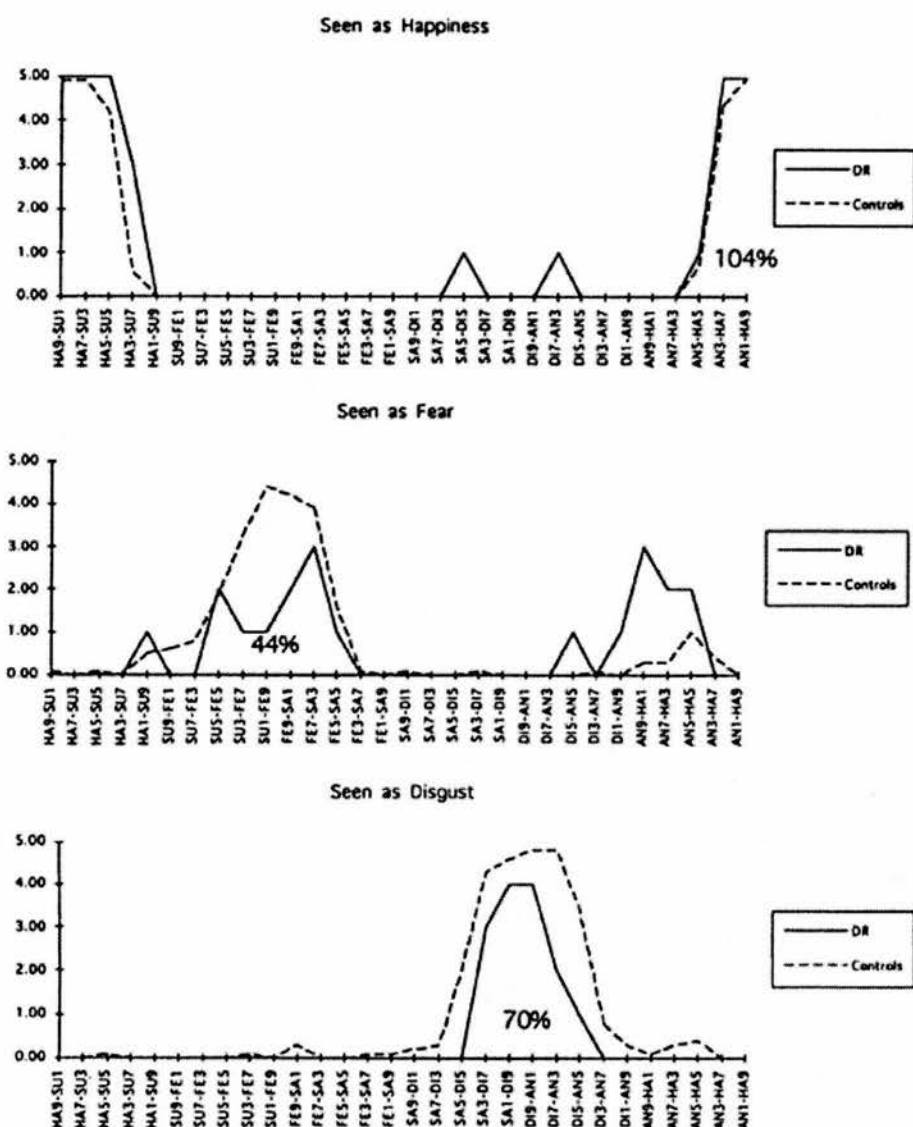


Figure 1.22 A sample of the data from Calder et al. 1996 using the Emotional Hexagon. The proportion of each emotion is presented along the x-axis e.g. HA9-SU1 relates to 9/10<sup>th</sup> happy 1/10<sup>th</sup> surprise. The y-axis relates to the number of times each stimulus was seen as looking happy (top graph), fearful (middle graph) or disgusted (bottom graph).

In the top graph, DR, an individual with bilateral amygdala lesions, can be seen to behave similarly to controls in terms of categorising the stimuli as happy. The lower 2 graphs, differences can be seen between the faces that DR and controls categorise as looking fearful and disgusted. (From Calder et al. 1996)

Later work using the imaging techniques of Positron Emission Topography (PET), has shown that the amygdala is indeed activated by viewing fearful facial expression. Morris *et al.* (1996) imaged subjects while they performed a gender discrimination task. The face stimuli presented during this task differed not only in gender but also in the amount of happiness and fear expressed. The stimuli were

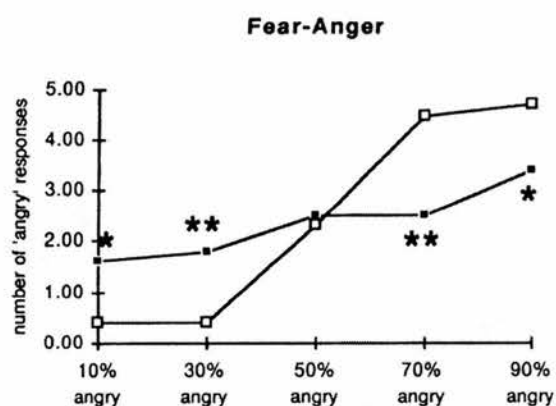
morphs between facial expressions of neutral-fear and between neutral-happy. To accentuate the amount of happy and fearful expressions the full blown expressions were warped into the shape produced by caricaturing<sup>13</sup> the differences between the neutral and the full blown expressions by 25%. Morris *et al.* found heightened activity in the left amygdala for expressions of fear, as compared to neutral, but not for expressions of happiness, as compared to neutral, suggesting that the amygdala has a particular specificity for facial expressions of fear. Later work, also using morphing techniques, has suggested that the amygdala is in addition sensitive to expressions of sadness, but not to expressions of anger (Blair, Morris, Frith, Perrett, & Dolan, 1999; Morris *et al.*, 1998) or to expressions of disgust (Phillips *et al.*, 1997).

Morphing techniques have also been used in the investigation of the effects of Huntington's disease on the perception of facial expressions. Huntington's disease is a neuro-degenerative disorder with widespread effects, including degeneration of the amygdala and basal ganglia which is the probable locus of damage causing the deficits in expression perception. Sprengelmeyer *et al.* (1996) used the emotional hexagon test and images from morphs between different pairs of expressions (sadness-happiness and fear-anger), a pair of familiar (famous) individuals (Cary Grant and Humphrey Bogart) and a pair of male and female gender averages. The emotional hexagon test was performed in the same way as described by Calder *et al.* (1996a) above. For the other test stimuli, subjects were presented with 10, 30, 50, 70 and 90% morphs and asked which of the original two descriptors, for example, sadness or fear, related best to each stimulus. There were no significant differences in scores between control subjects and individuals with Huntington's disease for the

---

<sup>13</sup> The method for caricaturing will be covered in the next section (page 55).

morph related to identity (Cary Grant-Humphrey Bogart), gender or sadness-happiness. However, for the morph fear-anger, Huntington's disease individuals' responses were flattened in comparison to controls as can be seen in Figure 1.23. Individuals with Huntington's disease performed poorly overall on the emotional



**Figure 1.23** Comparison of the number of 'angry responses' to a morph continuum running from fear to anger for control individuals (open squares) and Huntington's disease individuals (filled squares) (\*\* $p < 0.01$ ). (Figure from Sprengelmeyer et al. 1996).

hexagon test, being significantly poorer for each individual emotion, with the exception happiness, and particularly poor for recognition of disgust whose recognition was at chance.

Morphed facial stimuli have been

used in imaging studies aimed at the localization of the brain areas relevant to self-recognition (recognition of morphs containing the subject's face compared to

recognition of individuals who are highly familiar to the subject, Kircher et al., 2000; Kircher et al., 2001); to study the effects of psychoactive drugs on expression recognition (e.g. diazepam leads to a selective deficit of anger perception, Blair & Curran, 1999); and also to investigate depression (covered in Chapter 6), alcoholism (see appendix titled "Facial affect perception in alcoholics") and the developmental disorders including Williams and Down's syndromes (using animated morph expressions, Frigerio et al., 2001).

Individuals with developmental disorders like Williams and Down's syndromes manifest social problems that may be a result of difficulties in social interaction caused by abnormal comprehension of facially expressed emotion. Individuals with these syndromes may also have cognitive problems that make their

psychological testing difficult. To investigate these individuals' comprehension of facial expressions, a task was needed that was both simple and interesting to perform but required only a short concentration span. Facial expressions can be animated using a series of frames generated by morphing between an image of a neutral facial expression and a 'full-blown' facial expression to make such a task (Gagliardi et al., 2001). During the task, a face with a neutral expression appears on screen; the face is then animated to one of 5 facial expressions that the subject is asked to name. The intensity of expression is varied to avoid ceiling effects that are found for the identification of some expressions (in particular, expressions of happiness; for example see appendix titled "Facial affect perception in alcoholics"). Preliminary results with this test have shown that young individuals with Williams syndrome have similar accuracy rates for the identification of facial expressions to controls, and that individuals with Down's syndrome are substantially worse at recognising facial expressions than controls or individuals with Williams syndrome (Frigerio et al., 2001). For studies investigating emotion using animated stimuli, it may also be important to note that the rate of onset of facial expression, that is how fast the expression changes from a neutral to the full-blown expression, has effects on how facial expressions are understood.

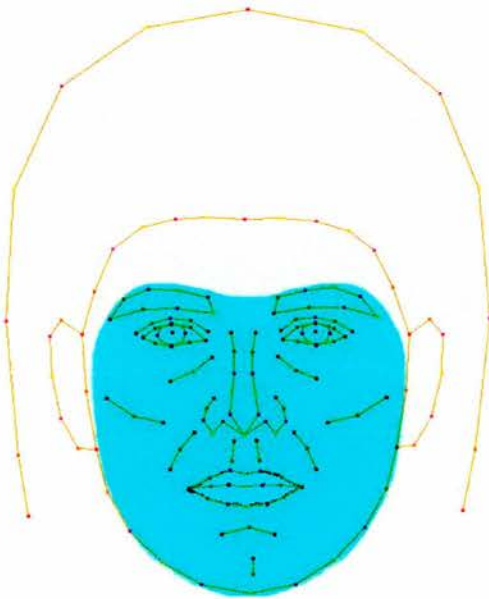
The effects of varying the rate of onset of facial expression sequences, created by morphing between neutral facial expressions and full-blown facial expressions on the perception of the facial expression, has been investigated by Kamachi, Bruce, Mukaida, Gyoba, Yoshikawa and Akamatsu (2001). The dynamic properties of facial expressions may be processed using separate neural pathways from static expressions (case study by Humphreys, Donnelly, & Riddoch, 1993). Kamachi *et al.* found that the rate of onset of facial expressions affected both subjects' accuracy in expression

identification and the magnitude of emotional feeling that was ascribed to the face. The effect depended upon the emotion being displayed (e.g. sadness was recognised more accurately with slow onset and happiness and surprise with fast onset). These effects not only suggest that it is important to standardise the frame rate when presenting animated facial expression stimuli, but also suggest a further area of study in groups suspected of abnormal emotional expression comprehension.



## ***Manipulation of individual features***

Most of the graphic manipulations presented in this introduction are applied to the whole face but occasionally it is useful to be able to manipulate just a part of the face. For example, Keating (2001) investigated how eye and lip size affects the perception of charisma<sup>14</sup>. This type of study is most suited to manipulation using a local transform of the area of interest. Ideally manipulations of features should be carried out in a manner that can be straightforwardly replicated on a number of face images, should be quantifiable, and easily visualized but should also produce high quality graphic output.



**Figure 1.24** An example template image used to mix the outer features from one image and the inner feature from another. Intensity values of the red channel are used for the mask. Areas to be taken from face B are shown in **turquoise** (RGB: 0,1,1) and areas to be taken from image A are shown in white (RGB: 1,1,1).

With these aims, a method was designed that made use of image warping and morphing techniques. This method (as used in Burt & Perrett, 1997 and Chapter 2) enables two face images, A and B to be mixed to make image C. A ‘template’ image is first prepared using a picture of an average face shape as a guide (Figure 1.24). Using a graphics package (such as Micrografx Picture Publisher 1987-1999), a template can easily be edited by adjusting the intensity values of pixels in the first

---

<sup>14</sup> In the study by Keating (2001) the features were manipulated by hand in a computer graphics program.

colour channel (Red). Areas of the template image that correspond to areas of the face to be taken from face A are set to 0 while areas to be taken from face B are set to 1. Intermediate intensity levels in the Red channel of the template image are taken from both image A and B in proportion to the intensity value. The template is normally blurred to produce a smooth cross dissolve in colouration and shape between images A and B.

Once a template image has been made, the manufacture of face image C is a two-stage process. In the first stage, the shape of face C is made. While in the second stage the colouration of image C is calculated. (If only a shape change is desired then only the first stage is used and image A is warped into the new shape, or, if only a colour change is required only the second stage is performed.) During the first stage the position of each feature point of image C is calculated. For each feature point the Red channel intensity at the position corresponding to the feature point in the template image is looked up. The feature point's position in face C is then calculated using the corresponding feature point values in the shape of A and B as:

$$C = ( \text{Intensity} * A ) + ( ( 1 - \text{Intensity} ) * B )$$

In the second stage, images A, B and the template image are warped into the new shape C. The above equation is then used to calculate the coloration of RGB pixels values in image C, based on corresponding values in the warped images A, B and the template image.

Using this method, the shape and the colouration values from the two original images are mixed in areas that where template intensity value is neither 1 nor 0. This means that the features of one face do not abruptly change to that of the second, but can be gradually faded helping to hide the junction between the images. The methodology was first used to mix between pairs of images to produce stimuli that



**Figure 1.25** Original images of young (20 years, top left) and old (50 years, bottom left) faces can be mixed together in shape and colouration (right).

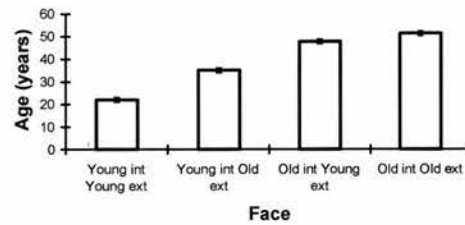
were used for the investigation of the lateralization of perception of facial attributes (see Burt & Perrett, 1997 and Chapter 2). The methodology has also been used to mix between pairs of young and old face images in an experiment to determine the relative effect of external features (hair and ears) and internal features in judgements of the age of an individual from their face.

A masking image was made up

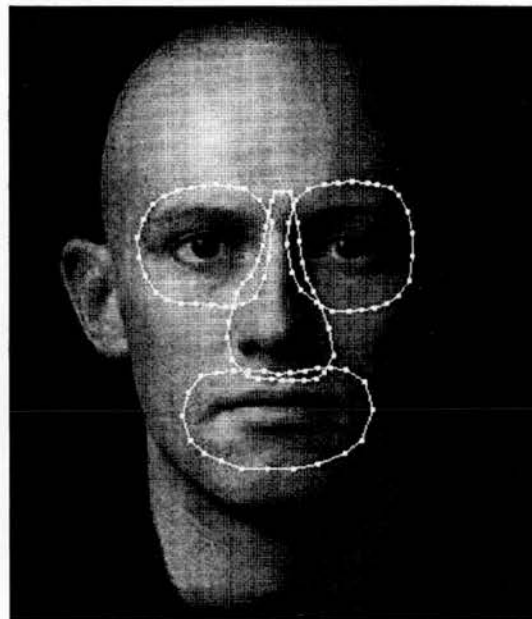
in which red channel values in the area corresponding to the internal features was set to 0 and the rest of the image set to 1. This masking image was then blurred. From a face image database six pairs of old and young face images were selected. The masking image was applied to each image pair to make an image with young internal features and old external features and a second image with old internal features and young internal features. Examples of these images are shown in Figure 1.25. Results from the use of this type of stimuli have shown that age perception depends both on internal and external facial features (Figure 1.26).

Other methods have also been used to mix the colouration values between pairs of faces. Steyvers (1999) presents a method in which parts of face A are bounded using groups of feature points (Figure 1.27). The colouration of a second image, face B, can then be mixed with the first image. Seams between the two images are rendered less visible by fading the image colouration between the two images at their boundary.

**Rating of age external vs internal features**



**Figure 1.26 The effect on perceived age of the old and young, outer and inner facial features.**



**Figure 1.27 Different features bounded by 'feature points' for use in the Steyvers method of feature swapping (From Steyvers, 1999).**

# Image and shape caricaturing

Morphing of both RGB colouration and the shape of objects in an image is a linear process. Caricaturing (Brennan, 1985; Rowland & Perrett, 1995) is the continuation of this linear process, which can, in theory, be carried out *ad infinitum*. However, in practice continuation of the linear changes of the shape and colouration of the image eventually result in artefacts and decreased realism, due to overlap of



Figure 1.28 Shape caricaturing expressions against different norms. Expressions were caricatured against either the average expression (left) or the neutral expression (right) by either -50% (half way to neutral), 0% (the original expression) or +50%. The lower 3 rows show the effects of caricaturing for happiness, surprise and fear. (Figure from Calder *et al.* 1997).

facial features and image colouration values topping or bottoming out<sup>15</sup>.

The process of caricaturing has been used to accentuate facial expressions.

Calder, Young, Rowland and Perrett (1997) caricatured the shape of facial expressions against either the average expressive face shape for an individual (making expressions of anger, disgust, fear, happiness, sadness and surprise) or the individual's neutral expression (Figure 1.28). Both forms of caricaturing were found to speed the identification of facial expressions to the same extent. (No significant effect was found for accuracy which may have been at ceiling in the uncaricatured images.) Calder *et al.* compared the two types of caricaturing since they suspected that the neutral face shape included information regarding the position of muscles when the face is in a neutral expression configuration. Therefore, caricaturing against neutral would simulate enhanced muscle action. By contrast the average face shape contains information regarding all of the facial expressions and caricaturing against the average shape makes the face more distinct from other emotional expressions. The two forms of caricaturing could thus, potentially differentiate between processes underlying expression recognition. The finding that the speeding of expression recognition is similar for both ways of caricaturing may be a reflection of the lack of sensitivity of reaction time measures with regard to picking out the small differences that can be seen to be present between the stimuli (Figure 1.28).

A later experiment by Calder *et al.* (2000) demonstrates that shape caricaturing has a linear effect on ratings of the intensity of the expression being caricatured, but that increased caricaturing also has the side effect of making the faces

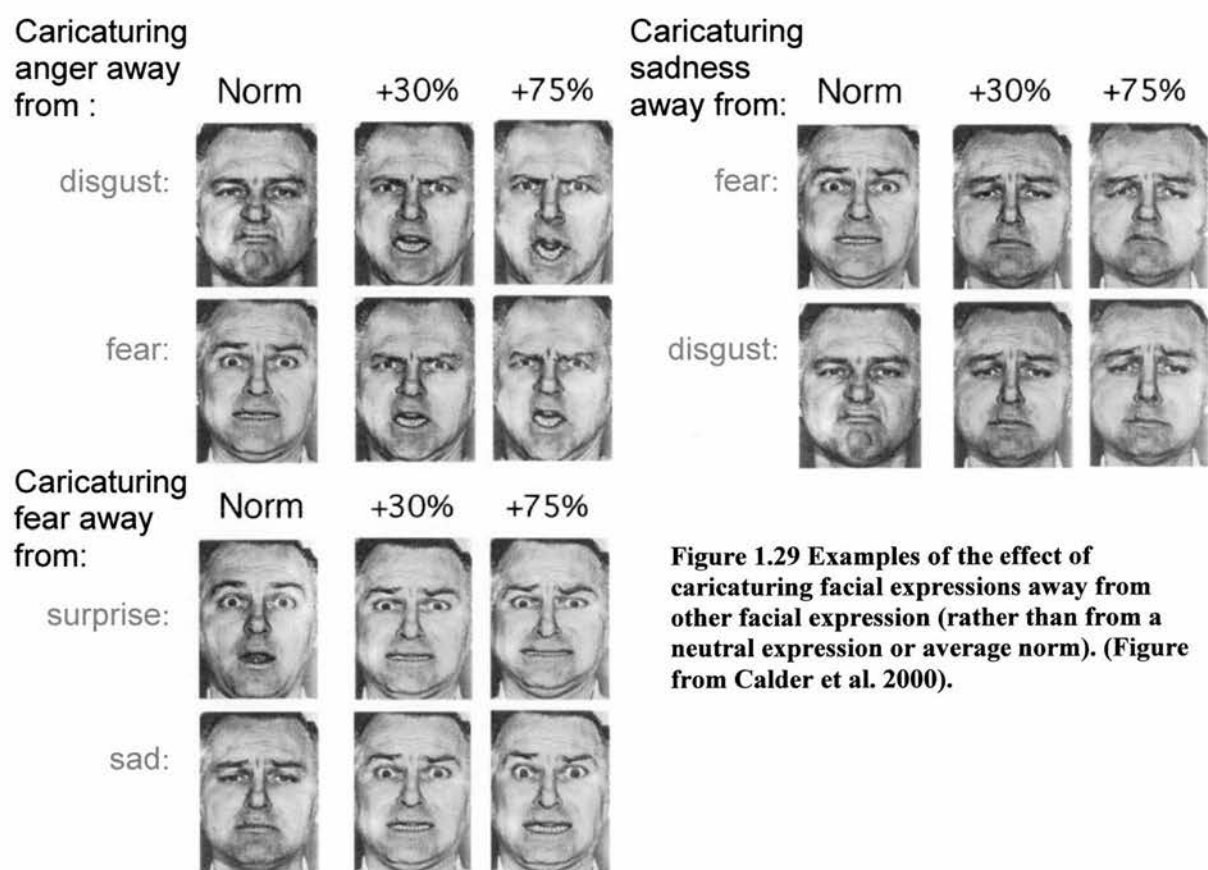
---

<sup>15</sup> Image intensity can only vary between set values, normally 0-255, for each RGB channel.

Caricaturing can lead to values out with this range. These cannot be properly represented by screen colours and are 'topped', or 'bottomed' out.



look less ‘face like’ to subjects. Calder *et al.*, 2000, also extended their 1997 study by caricaturing against, not only the neutral expression configuration and the average expression configuration, but also other facial expressions. They found that caricaturing against other expressions (Figure 1.29) had the effect of increasing ratings of the emotional intensity. Recognisability of the caricatures in terms of accuracy or effect on reaction time was not tested. Calder *et al.* (2000) interpreted

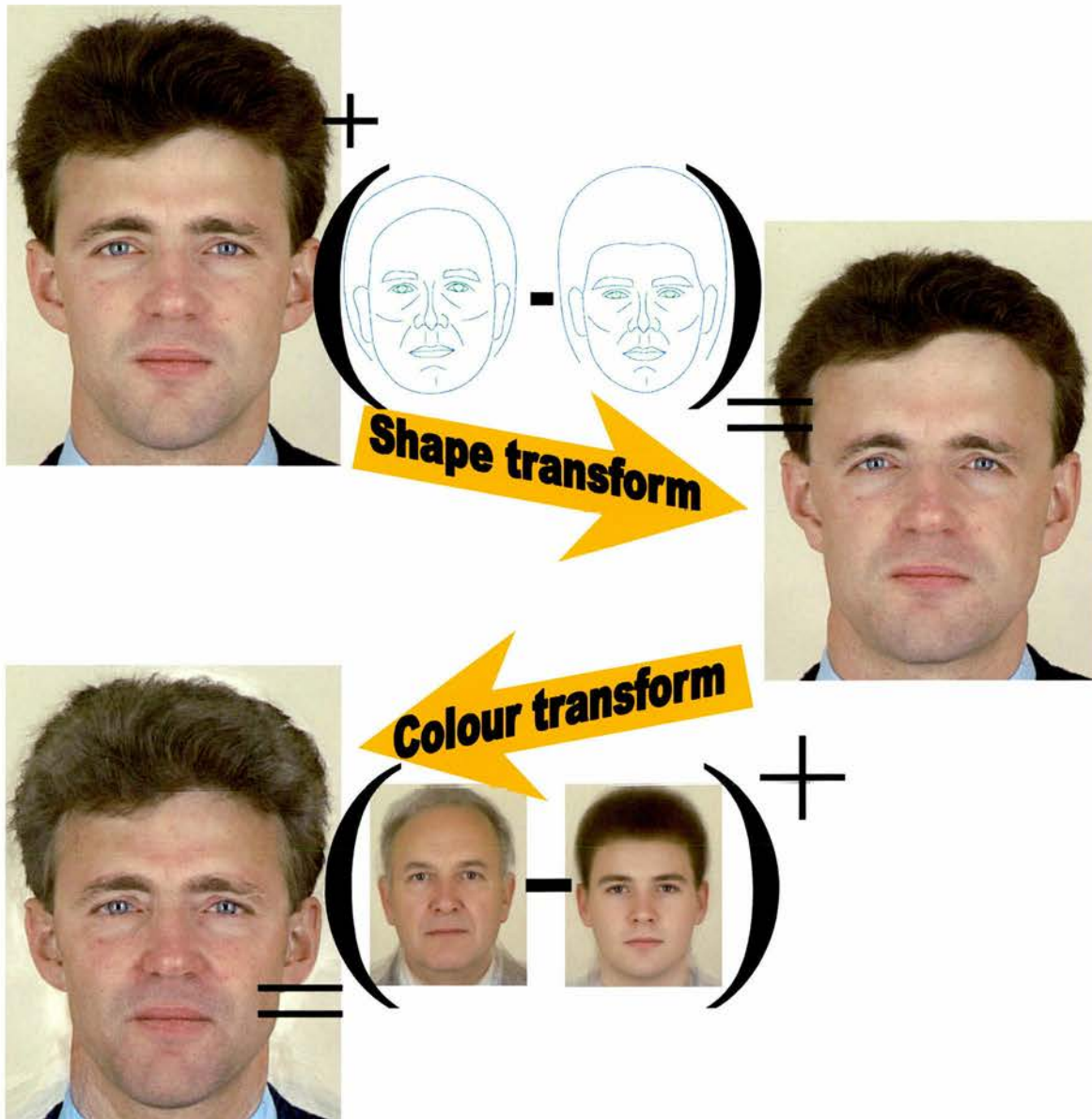


their results in terms of an exemplar based model of emotions for the recognition of expressions. Alternatively, the results might be explained by higher levels of caricature resulting in faces that appeared more intense in general because of greater levels of distortion. Subjects could interpret the distortions present in the face images as a sign of intense feeling rather than as indicative of the particular type of emotion present.

The combination of both image and shape caricaturing has also been widely used to study facial expression perception. Later in the thesis caricaturing is used to accentuate the colouration differences between average images manufactured on the basis of ratings of skin health (Chapter 2). The difference between a pair of images can also be applied to alter a third image with a technique known as transforming. This technique is covered in the next section.

## ***Image and shape transforms***

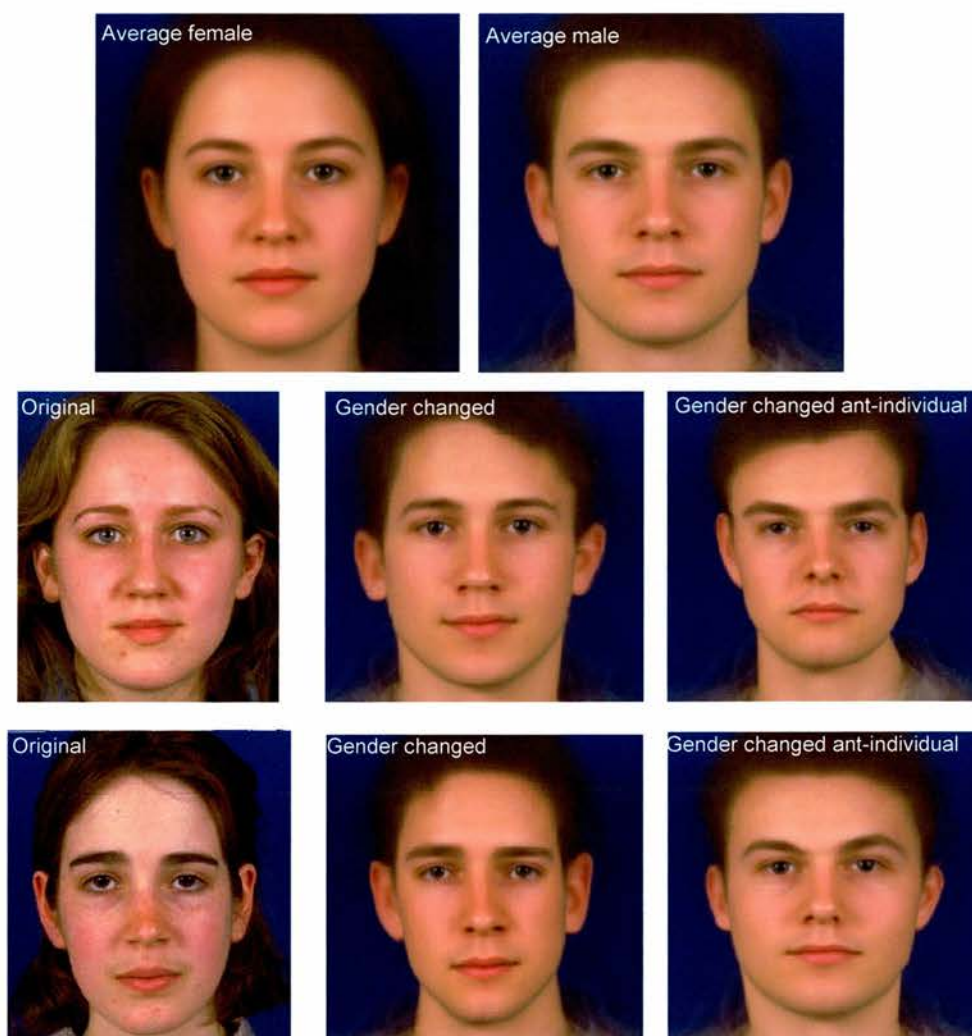
Image and shape transforms can be regarded as an extension of the caricaturing technique. Whereas in caricaturing the difference in shape and colour occurring between one image and another is exaggerated, in transforms the difference between two images is applied to a third to change the shape and colour of objects



**Figure 1.30** The difference between the average face shape of a group of old adults and the average face shape of a group of young adults can be applied as a shape transform to the face of an individual (top left) to give a new older shaped face (right). The difference in colouration between the old and young average images can then be applied as a colour transform to the face to give an older shape and colour face (bottom left).

within it, as shown in Figure 1.30 (Burt & Perrett, 1995; Rowland & Perrett, 1995; Yamada, Chiba, Tsuda, Maiya, & Harashima, 1992).

One of the first studies to use a shape transforming technique was performed by Penton-Voak, Perrett and Pierce (1999b). The study investigated the apparent paradox that individuals seem to have a tendency to look like their partners (as shown by studies demonstrating that when asked pair up couples from separate photos subjects show greater than chance matching of the real couples' photographs e.g.



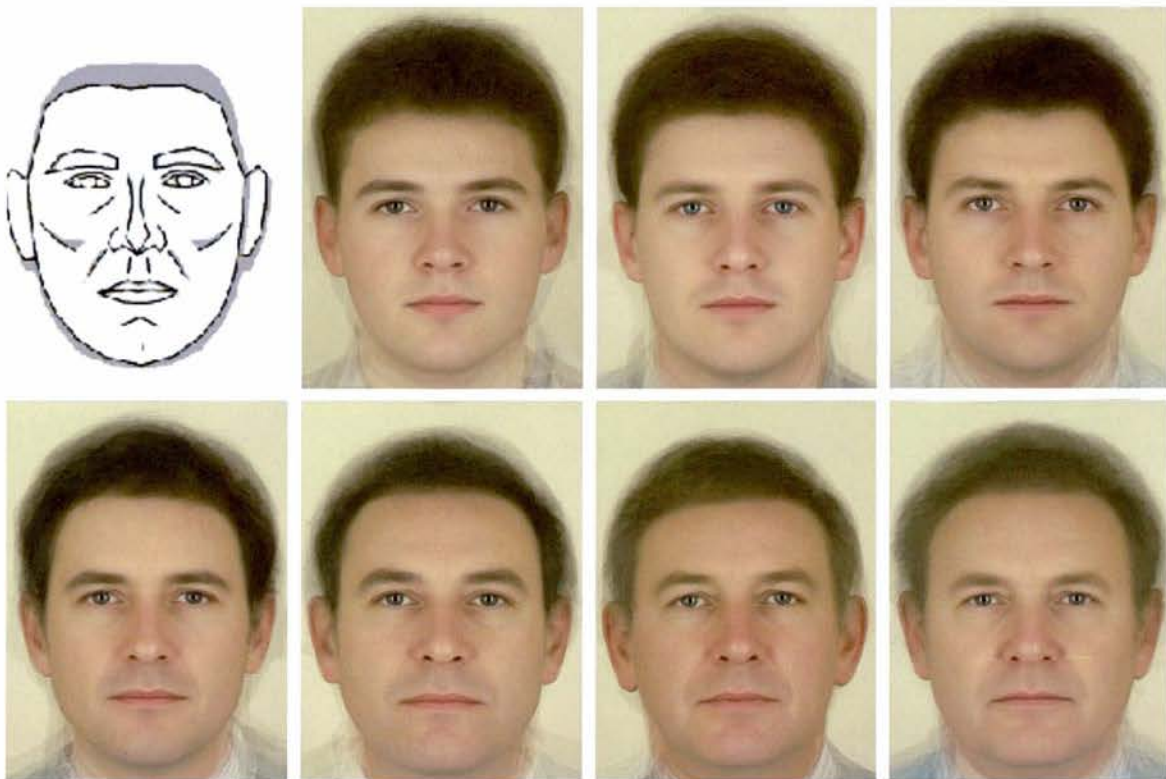
**Figure 1.31** The difference in shape between the female average and the male average (in the top row) was applied to the shapes of the individual faces (lower left) to change their gender. The male average texture was then warped into these new face shapes to make opposite sex faces (lower middle). Graphics hardware was used to create continuum that ranged in shape from the opposite sex faces (lower middle) to the average shape and then to 100% 'anti-individual' (lower right).



Griffiths & Kunz, 1973) even though inbreeding is often thought to be detrimental (e.g. Bateman, 1983). Penton-Voak *et al.* (1999b) investigated whether individuals find novel, but similarly shaped, opposite sex faces more attractive than dissimilarly shaped faces. Subjects were presented with opposite sex facial stimuli on screen for manipulation along a continuum (by moving the mouse from left to right). The continuum ranged from shapes that were more like the subject than the average face shape, to the average face shape and then to face shapes that were less like the subject than the average face shape, the anti-individual (Figure 1.31). No effect of subjects' own face shape was found on the selection of the optimally attractive point in the continuum. This null result may be due to the small scale of self-similar preferences in comparison to the scale of preference for the most average face shape. The preference for average shape in the experiment may be the result of the methods used by Penton-Voak *et al.*, since the average face was presented multiple times during a test session (possibly inducing a mere exposure effect) and the image texture that was used corresponded to the average face shape and was therefore less aberrant in the average shape.

Image and shape transforms have also been used to investigate the perception of age in faces. Burt and Perrett (1995) took the difference in shape and colour between young and old average images and applied this information to individual face images. Previous studies of ageing have tended to concentrate on a model of skull growth named cardioid strain (Pittenger & Shaw, 1975). The changes induced by cardioid strain are simplifications of the changes in the shape of the skull that take place during growth, but are unlike the changes that occur with ageing in adulthood as represented in Figure 1.32. This is because most of the shape changes in adulthood are due to skin elasticity changes (wrinkles and sagging) and soft tissue growth (in the

ears and nose). The shape and colour differences between the average faces produced from two different age groups (in this case, of average chronological age 27 and 52) can be applied to individual faces to increase the perceived age of the face (Figure 1.30). The changes in perceived age that were induced by the age transforms were not



**Figure 1.32** The difference between young, average face shapes of the 25-29 (dark lines), and older male faces, 50-54 age brackets, is illustrated by the shaded area. Each of the blend images contain the average colour and shape information from 15 individual faces within the same 5 year age bracket: 20-24, 25-29, 30-34, 35-39, 40-44, 45-49 and 50-54. (Figure from Burt and Perrett, 1995)

as great as might have been expected. For young faces (aged 27) the shape and colour transform added 12 years to the face rather than the 25 years that might have been expected given the average chronological ages of the blends on which the transform was based. This is approximately half the age change that might be expected and is probably due to the loss of textural information, like skin wrinkles, in the blends on which the transforms were based. Subject ratings of the age of transformed faces also suggest that the colour and shape changes are non-linearly related to perceived age,



since applying the same ageing transform to older faces had less effect on perceived age than applying the transform to young faces.

The next section will examine recent computer graphic methods related to image texture aimed at addressing the loss of textural information when images are blended together and transformed.

## ***Texture processing***

Blending together different images results in an image that is smoother in texture than the originals, and is therefore atypical. One example of this effect is that blends of faces tend to look younger than their components. Thus, subjects perceive the face images on the left hand side of Figure 1.33 as being up to 20 years younger (Burt & Perrett, 1995) than the actual chronological age of the individual face images

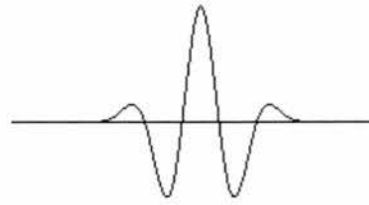


**Figure 1.33 Female face image blends: normal versions (left) and textured versions (right). For young faces (ages between 20-29; top) and for older faces (aged between 50-59; bottom).**

from which the blend images were produced. Tiddeman, Burt and Perrett, D. (2001a) used a technique based on wavelets to reconstruct the textural information that is lost

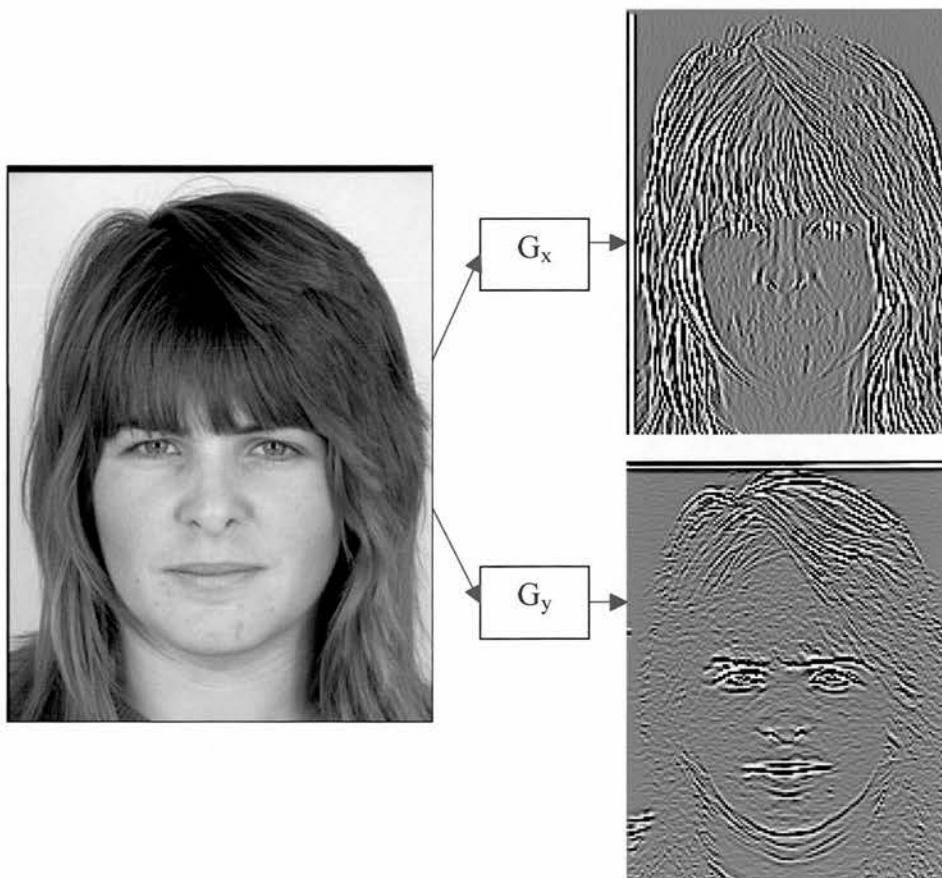
during image blending. An example of the shape of a wavelet is illustrated in Figure 1.34.

Wavelet like filters are often used in the detection of edges in computer graphics and are an approximation of the type of filtering that is employed early in the visual system (Daugman, 1980). The visual system uses filters of many different orientations. Constraints, resulting



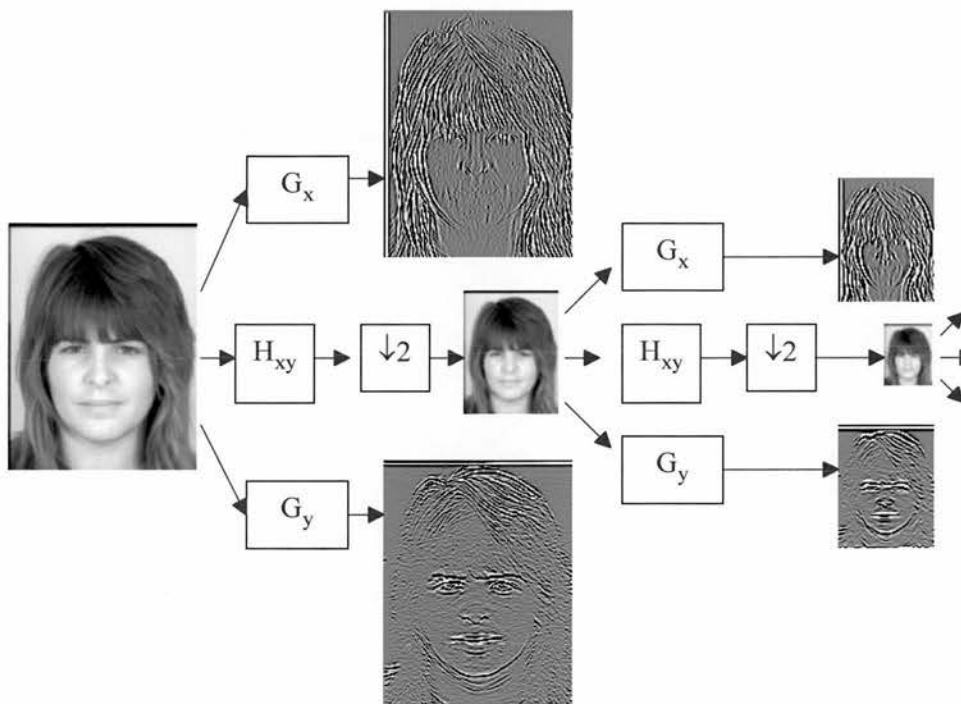
**Figure 1.34 A wavelet as used in texture processing (Figure from Tiddeman et al. 2001a).**

from computer processing power and information storage, led Tiddeman *et al.* to use two orientations (horizontal and vertical). The effect of passing a vertical wavelet filter over an image can be seen in the top image of Figure 1.35. The amplitude of high spatial frequency, horizontal contours within the image is quantified by scanning the image with the same filter orientated horizontally (Figure 1.35, bottom). After



**Figure 1.35 The original image on the left is filtered with a wavelet filter horizontally to produce the image on the upper right and vertically to produce the image on the lower right. (Figure from Tiddeman et al. 2001a)**

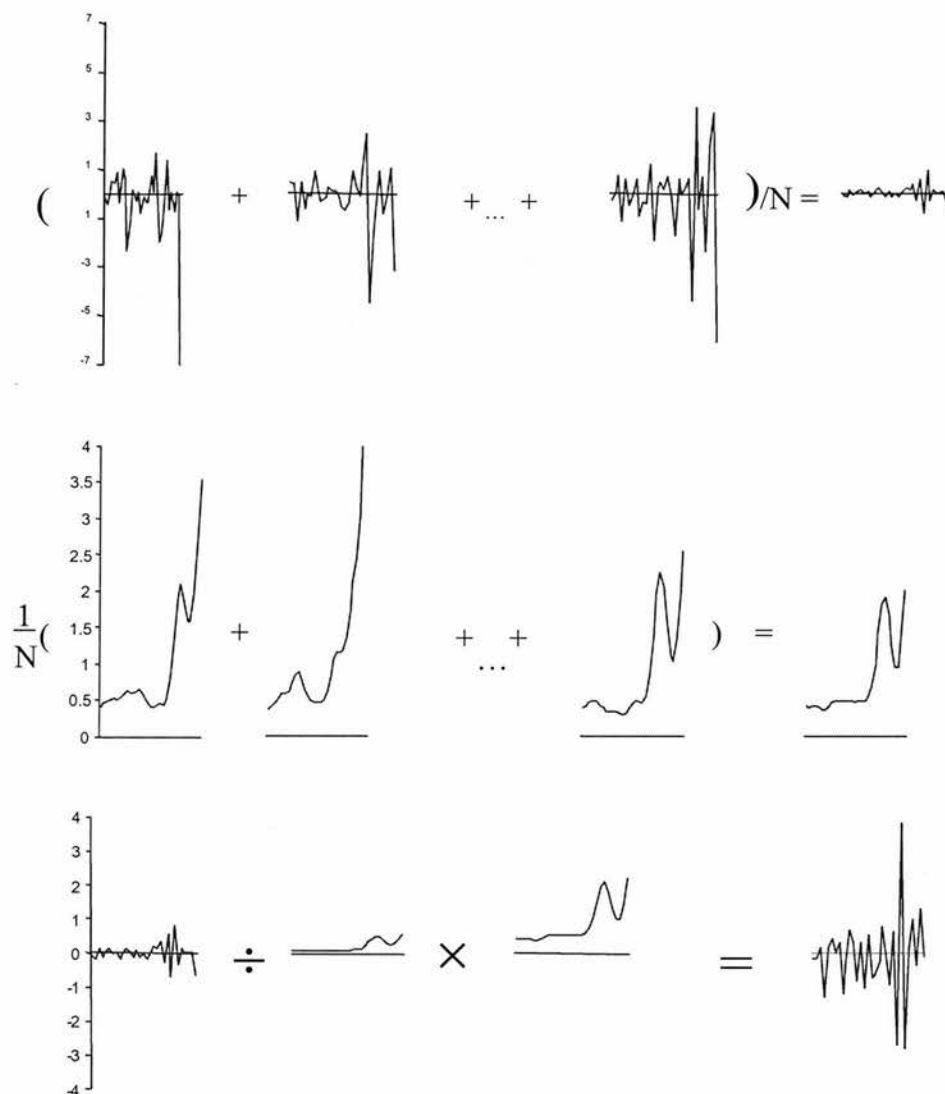
being scanned horizontally and vertically the image is resized to a quarter of its original size and the process is repeated (Figure 1.36). The same absolute size of wavelet filter is applied vertically and horizontally to filter the smaller image. This has the effect of halving the spatial frequency of contours that are detected. This process of resizing and filtering is then applied repeatedly to produce a wavelet decomposition pyramid for the starting image. This process is reversible. The image decomposition pyramid can be converted back, ‘collapsed’, into a very good approximation of the image from which it was created.



**Figure 1.36** The process shown in Figure 1.35 is repeated. The size of the image is halved to produce a new image over which the wavelet filters are run. This produces a wavelet decomposition pyramid. (Figure from Tiddeman et al., 2001a)

The averaging together of images causes textural details to be smoothed. The same effect appears if wavelet decompositions are averaged together. This smoothing is represented in the first row of Figure 1.37. It can be seen that the wavelet amplitude has been reduced. The average magnitude of the wavelet decompositions of component images can be calculated (second row Figure 1.37). The average wavelet decomposition is then rescaled to reflect the individual wavelet magnitudes (Figure

1.37, bottom row). The resulting image decomposition pyramid can then be collapsed to produce a textured average image.



**Figure 1.37** Wavelet magnitudes from a set of images (image 1 plus image 2 ... plus image N) can be averaged together (top row), but this average (right) is not representative of the set.

To make a more representative average the average unsigned wavelet magnitude is calculated (middle line).

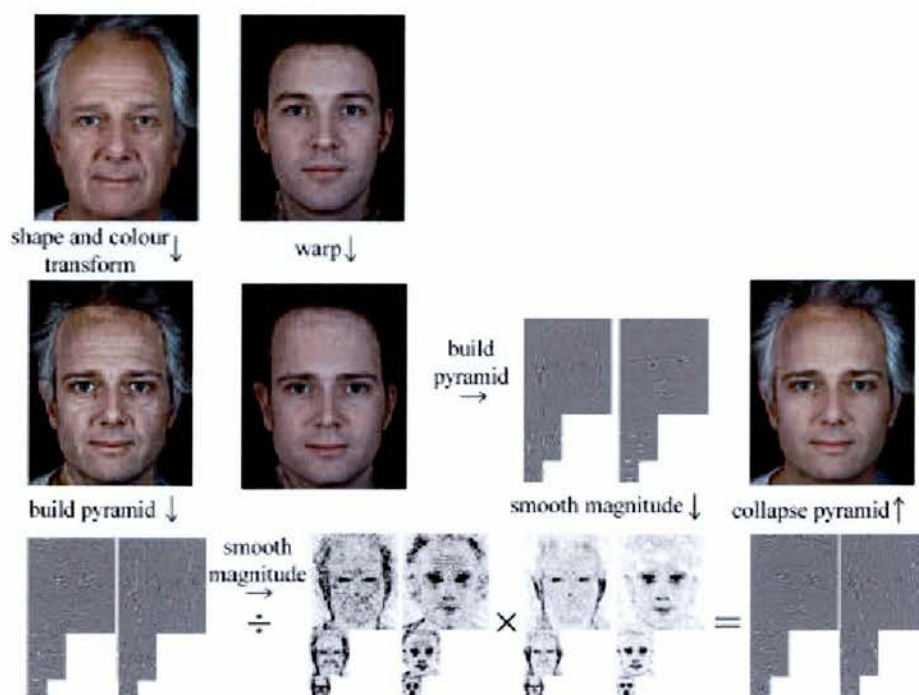
The average wavelet magnitudes (bottom right) is then rescaled. This is performed by dividing it by the smoothed average wavelet magnitude and then multiplying the result by the smooth unsigned wavelet magnitude to arrive at a better representation of the original images (bottom right). (From Tiddeman et al., 2001a).

Examples of blend average images manufactured with the benefits of this method can be seen in the right column of Figure 1.33. These ‘textured’ images look more realistic and are perceived to be the average chronological age of the group of faces from which they were composed (Tiddeman et al., 2001a).

Image transforms also benefit from the use of texture processing based on wavelets. One of the ways of using texture in transforms is to apply a transform in shape and colour to an image, and then to adjust the wavelet magnitude values, related to the manufactured image’s texture, to better reflect those of the average to which the



image is being transformed. This method is used by Tiddeman *et al.* (2001a) and is illustrated in Figure 1.38.



**Figure 1.38 Shape and colour age transform using texture adjustment.** In this method, the texture magnitude from the average young face is used to replace the texture of the face after it has been colour and shape transformed to make it appear younger. (Figure from Tiddeman *et al.*, 2001a)

The effects of age transforms using texture magnitude adjustment to increase and decrease the age of faces can be seen in Figure 1.39. These texture transforms have the effect of ageing the face by the appropriate amount given the perceived age difference between the two groups of faces that were used to produce the averages used in the transforms. These texture transforms also have the effect of rejuvenating the face by around 45% of the expected amount (Tiddeman *et al.*, 2001a). The use of texture adjustment overcomes some of the problems related to idiosyncratic cues to age that are of low frequency in the population from which the average is derived. For example, one individual may have numerous wrinkles on their forehead, whereas another might have liver spots on the upper parts of their cheeks. Adjusting the textural information can reset many of these cues if they appear in the luminance domain.





**Figure 1.39 Effect of adjusting texture when performing transforms. Four original face images (left) are transformed to rejuvenate (a,c) and age (b,d) them in colour and shape (centre of each row), and with the addition of wavelet texture adjustment (right of each row).**

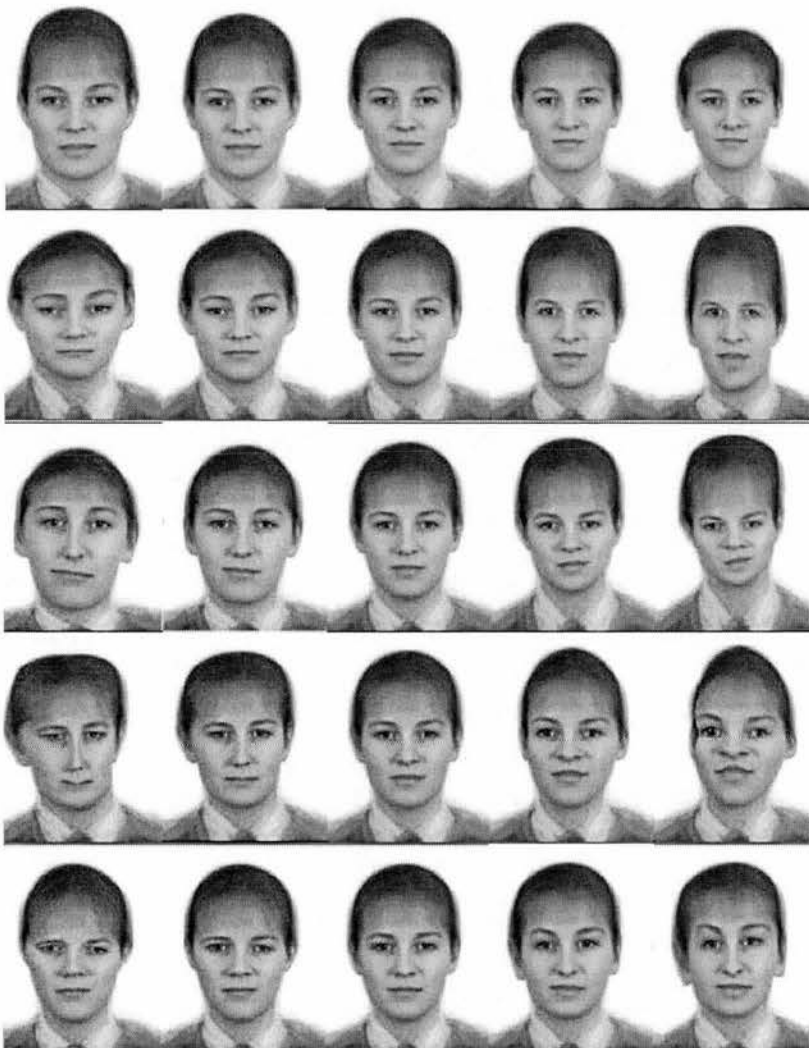
Texture processing was used in the generation of the anti-emotion stimuli described in Chapter 6 and in the morphed expression sequences used in the experiment with alcoholic subjects presented in the appendix titled “Facial affect perception in alcoholics” at the end of this thesis. The method used in these experiments is slightly different from the colour and shape transform followed by the texture adjustment method outlined above (Tiddeman et al., 2001a). In the experiments presented in this thesis that use texture transforms, the wavelet decompositions for each of the images used undergo a transform process akin to that

used for shape and colour transforms. This method produces less rejuvenation of the face images than the method given by Tiddeman *et al.* (2001a).

This thesis does not make use of all available computer graphics methods. Three areas of methodology that have not been used, but are applicable to the investigation of face perception, are methods based on Principal Component Analysis (PCA), on evolutionary algorithms for searching face space and on 3D face capture and manipulation. Since these methods are not used later in the thesis, they will only be reviewed briefly here.

## ***PCA analysis of facial variation***

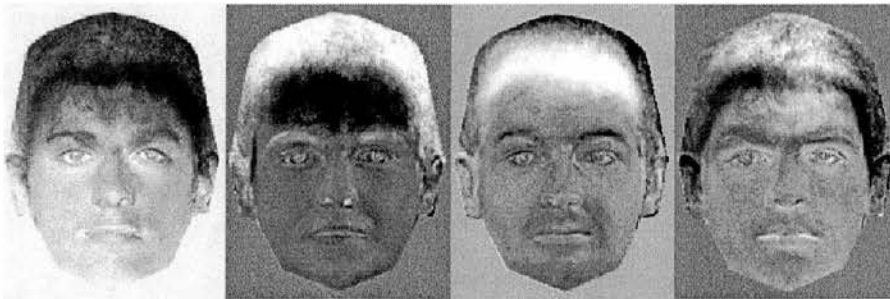
Principal Component Analysis (PCA) is a method of factor analysis in which a data set is transformed into a new set of orthogonal linear axes. The first of these axes or components accounts for the maximum amount of variation in the original data set. Later components account for progressively less of the variation from the data and are orthogonal (uncorrelated) to the first and subsequent components.



**Figure 1.41** The first 5 components from an analysis of shape variation in a set of faces. Face images are first delineated and a PCA performed on their shape data. The average face shape (centre of each row) is manipulated by the first principle component (top row), and next 4 principle components (subsequent rows) of the set of faces' PCA face space. The average faces' colouration information is used for the illustration (Figure from Hancock 2000)

Often PCA is used as a method of data reduction by selecting only the first few components with Eigenvalues greater than 1 for further analysis. Other times, a predetermined number of components or the full set of components is used (e.g. Deffenbacher, Vetter, Johanson, & O'Toole, 1998; O'Toole, Vetter, Troje, & Bulthoff,

1997a; O'Toole, Vetter, Volz, & Salter, 1997b). The changes related to the different components can be applied to the average face for visualization. For example, Figure 1.41 shows the shape changes related to the first five components from an analysis of shape variation and Figure 1.42 shows the effect of the first four components of image intensity information from the analysis of a set of male face images that have been warped into the same shape.



**Figure 1.42 Effects of the first 4 image intensity components applied to the average face. (Figure from Hancock, 1998)**

PCA analysis of the image intensity and shape from photographs have, following Turk and Pentland (1991), been used widely in face perception research for the evolution of faces (e.g. Hancock, 2000), for the examination of theories of face space (e.g. Hancock, Bruce, & Burton, 1998) and for both caricaturing and manipulating faces in 3D (e.g. Blanz & Vetter, 1999; Leopold, O'Toole, Vetter, & Blanz, 2001; Vetter, Jones, & Poggio, 1997).

# Genetic algorithms as a method of searching face space

Genetic algorithms can be used in the search to develop solutions to a diverse range of problems (Holland, 1992). Genetic algorithms are based on the notion of making ‘fit’, or preferred, individuals more likely to contribute ‘genes’ to the next generation.

There are different ways to specify the ‘genes’ for evolution. In the first method illustrated here the genes are interpreted, based on their position in the ‘genome’. Genes representing values for different parameters that govern the design

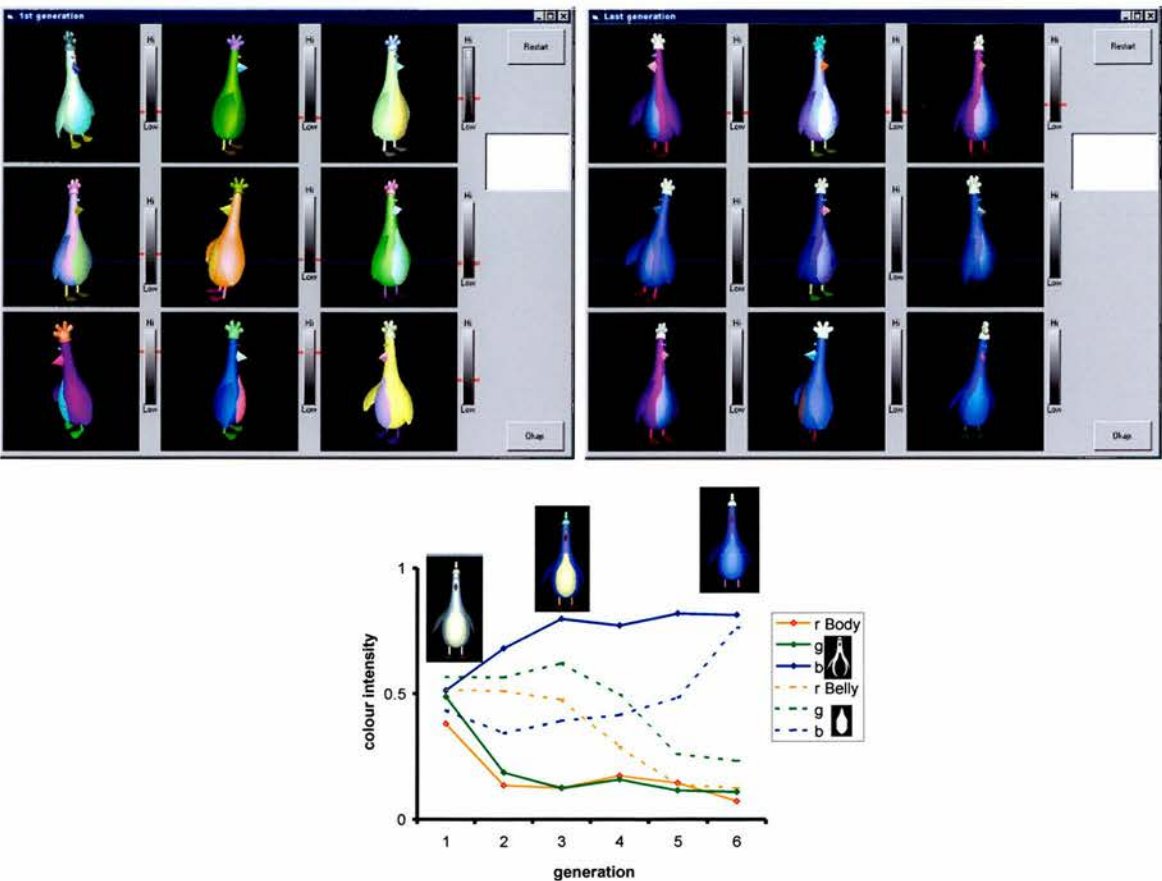
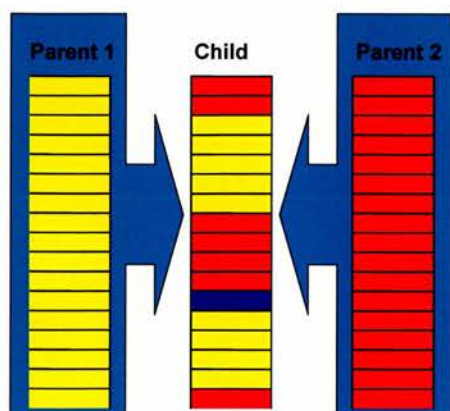


Figure 1.43 Evolving a blue chicken. A simple demonstration of an evolutionary algorithm and interface for optimising the colouration of a model based on subject ratings. An initial population of chickens, with random colours (left), is evolved to the 6<sup>th</sup> generation (right) of the subject expressing a preference for blue ‘chickens’. The graph represents the average red, green and blue colour values for the individuals in successive generations. (The 3D model of the character "Feathers" McGraw dressed as a chicken from the “The Wrong Trousers” © Aardman Animations.) (From unpublished work by Burt, Xiao, & Perrett.)



of an object have to be specified. The specification of genes is often called parameterisation. Parameterisation involves matching up ‘genes’ to the different parameters governing the design of the object to be evolved. For example, the colouration of the 3D model in Figure 1.43 can be parameterised with each component (the beak, body, belly, feet and glove) having 3 values, one for each of red, green and blue, each of which may vary between 0 and 1. The colour values for



**Figure 1.44 Schematic of genomes of parents and child. A first randomly sized block of bits was taken from one parent, parent 2 (red) in this case. Subsequent randomly sized blocks of genes are from alternating parents (parent 1 coloured yellow in the first case). Until the child has a full complement of genes. A ‘genetic mutation’ is represented in blue.**

the five components could be, for example, the 1<sup>st</sup> to the 15<sup>th</sup> values or genes respectively in the genome<sup>16</sup>. A genetic algorithm can then be used to evolve the object’s colouration. In the example illustrated in Figure 1.43, each colour value started as a random floating-point number

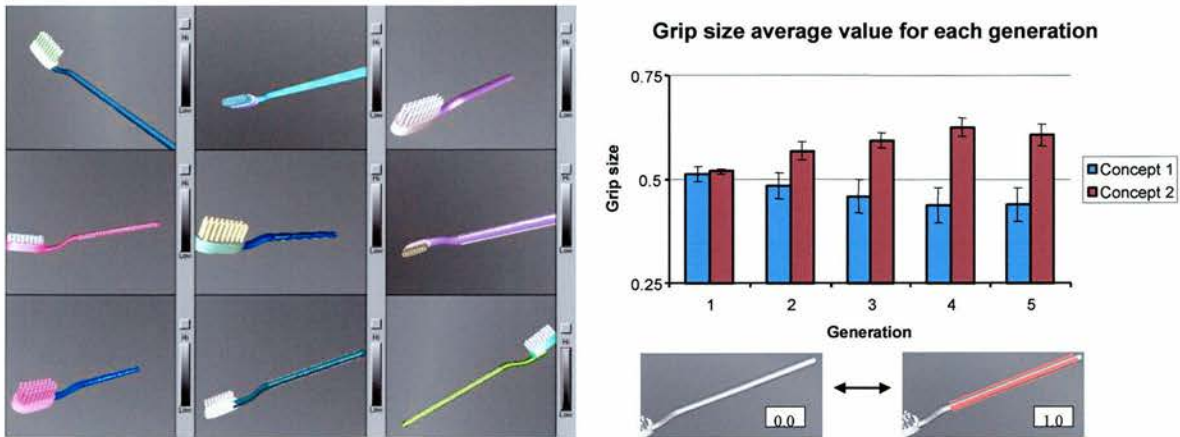
between 0 and 1, hence the objects are multicoloured. Preference for individual chickens was rated on a scale between 0 and 1. To manufacture the genetic makeup of a new individual, 2 parents were chosen from the

previous generation. The probability of an individual from the previous generation being chosen as a parent was calculated as the squared preference score for that individual, divided by the sum of squared preference scores for all of the individuals in the generation. The genes of the 2 selected parents were represented as a series of bits. The “child’s” genome was built from the randomly sized blocks of bits from

<sup>16</sup> In the example in Figure 1.43, other colour parameters governing the object were also evolved (e.g. reflected colour).



corresponding parts of the genomes of both parents (Figure 1.44). There was a small random probability of mutation (bits being flipped).

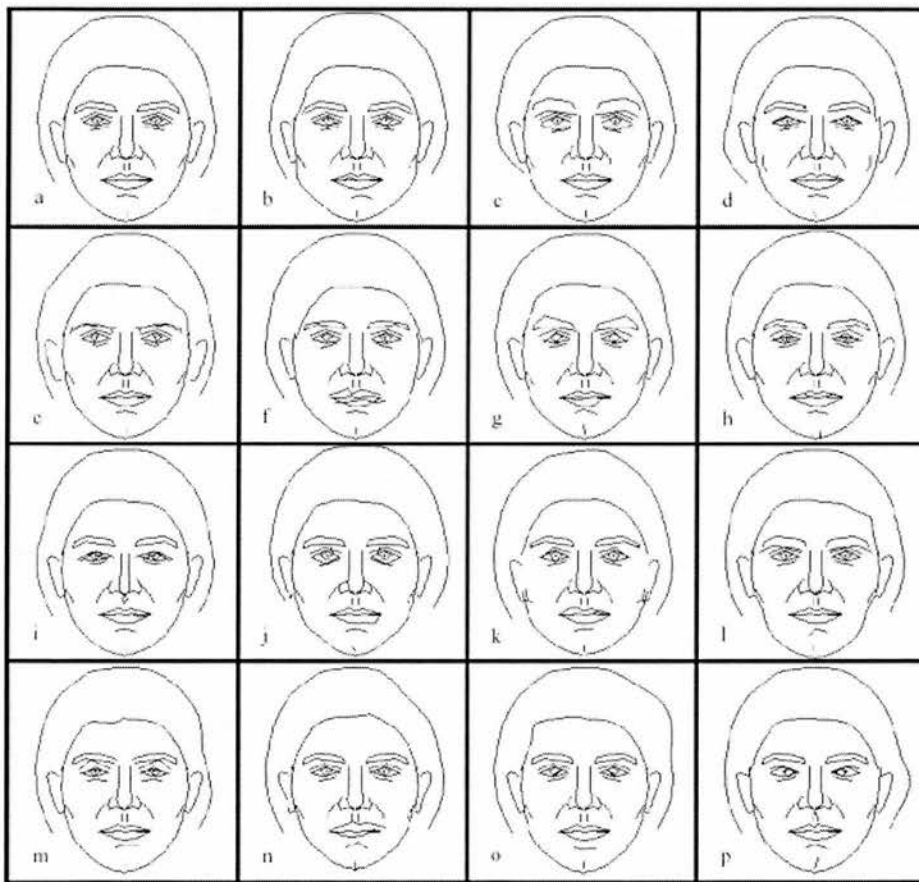


**Figure 1.45 Evolution of product design.** An example of toothbrush evolution for two different consumer concepts. Using a computer interface, subjects were able to visualize and rate 3D toothbrushes. These were evolved based on the subject's ratings for being good at concept 1 or concept 2. Results of evolution can be expressed in terms of change in features. For example, the graph right shows divergent change in grip size with generation when subjects were asked to evolve toothbrushes for the different concepts. (From unpublished work by Burt, Xiao, & Perrett.)

In Figure 1.43 the user preferred bluer birds and rated them more highly. The increase in the blueness of colour over successive generations can be seen for a selection of individuals after 6 generations and for the average RGB colouration of the individuals' body and belly areas for the first 6 generations. In a similar manner, 3D shape descriptions of objects can be parameterised to enable optimisation of product design using evolution based on ratings by non-experts for different concepts (e.g. toothbrushes Figure 1.45 or shampoo bottles, see Rowland, 1998). This type of method is constrained in that it only allows evolution of the pre-selected parameters.

A more unconstrained type of evolution is possible by using a genotype in which different pieces of 'genetic code' can be added. For example, Baker and Seltzer (1994) evolved line drawings. In the genotype of an individual drawing, each gene describes a line in full and so individuals can have differing numbers of genes. This lack of constraint means that, in theory, the line drawing could depict anything. The problem with this type of genetic model, with its lack of constraint, is that most of the

individuals generated will not resemble what the user wants to evolve. Most of the changes caused by the genetic algorithm randomly altering or adding lines are unlikely to make realistic looking faces. As can be seen from Figure 1.46, this method can be used to evolve drawings of faces, but most researchers have chosen to use more constrained methods in an effort to keep the individual faces produced by the genetic algorithm as face-like as possible<sup>17</sup>.



**Figure 1.46** A selection of examples of the effect of single mutations each starting with the average face (a). Although some of the mutations result in realistic face images many can be seen not to do so. (From Baker and Seltzer, 1994)

One method of constraint is to use databases from which the face parts are selected by the genetic algorithm. Caldwell and Johnston used this method to evolve

---

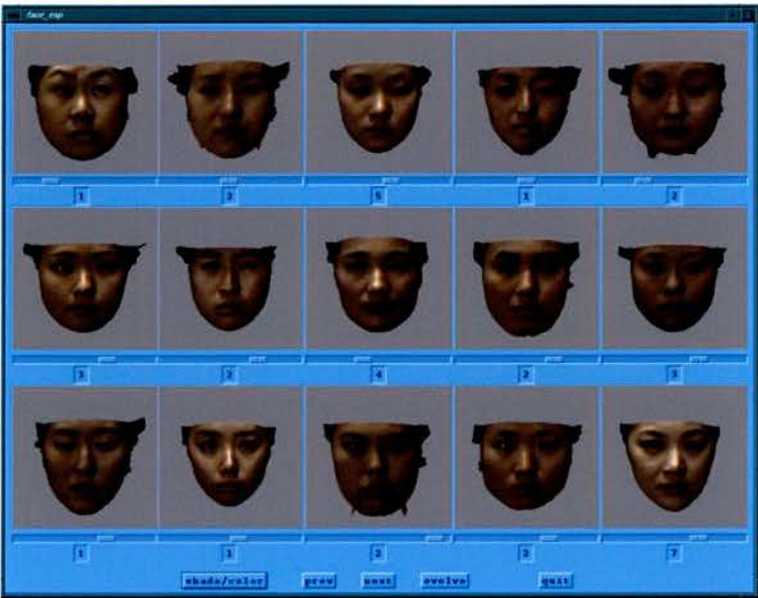
<sup>17</sup> Genetic algorithms are sometimes used in processes where a preference value can be quickly calculated by another algorithm. In comparison, humans are relatively slow at providing feedback on the individual faces and tire after rating a small number of individuals. Therefore for humans to use genetic algorithms, the method must be optimised to evolve as quickly as possible.

faces (Caldwell & Johnston, 1991; Johnston, 1994; Johnston & Franklin, 1993; Johnston et al., 2001). The method uses a set of databases of different face parts (a mouth database, a pair of eyes database, a hairline database etc.). Each face part in a database is numbered and a genetic algorithm is used to choose an example of each type of face part that is then pasted together to form a face. The exact position that individual features are pasted in is also coded by genes and evolved. A disadvantage of this method is that the position of exemplars in the database does not necessarily relate to the qualities of the individual exemplars. So, for example subsequently numbered noses will on average bear the same resemblance to one another as noses chosen at random. This has the effect that breeding by combining parents with, say, nose number 13 and nose number 3 will result in, for example, nose number 7, a nose that will on average be as similar to the 'parental' noses as any other nose selected at random. Thus, mating individuals with nose 13 and 3 does not produce a nose that is a combination of individuals and so information related to the subject's preferences is lost. This problem is resolved by the use of PCA for the parameterisation of face space (Hancock & Frowd, 1999).

Genetic algorithms have been used in combination with PCA to evolve faces as an alternative to using an identikit kit for face identification by witnesses (Hancock, 2000; Hancock & Frowd, 1999) and to evolve 3D faces for attractiveness research (Figure 1.47, Isono, Oda, & Akamatsu, 1997; Rowland, 1998). Rowland presented Caucasian subjects with evolved sets of 3D face shapes for rating of attractiveness with the aim of evolving more attractive faces. Rowland found the evolution achieved positive results; subjects preferred the individual faces that they had manufactured from later generations. It was found that, in comparison to generation 1, faces in generation 6 were a greater distance from average (in terms of

the average magnitude of Eigenvector values), a result that is interpreted in terms of a preference for non-average faces and therefore, against a strong version of the averageness hypothesis (Langlois & Roggman, 1990).

The use of PCA on sets of face data results in the largest changes within the



**Figure 1.47** The interface for rating and viewing 3D faces for evolution and a sample of 3D faces. (From Rowland, 1998).

group of face being reflected in the first components. Evolution may be speeded up by using only the more compact description of face space, described by the first few Principal components, with only a relatively small loss in the amount of variation in the individual faces that can be produced. Hancock and Frowd used this compact description of face space in research aimed at evolving photo-fit faces based on witness-guided evolution (Figure 1.48 Hancock, 2000; Hancock & Frowd, 1999). Using the first 18 Principal components from a PCA conducted upon a set of 35 faces (giving a total of 34 components), subjects used an interface (similar to that shown in Figure 1.48) to select the faces that they felt were most similar to a target face. To evaluate the success of evolution, an algorithm-based index of the difference between the evolved faces and the target face (mean-squared error of corresponding pixel intensities) was used. It was found that with successive generations the faces became



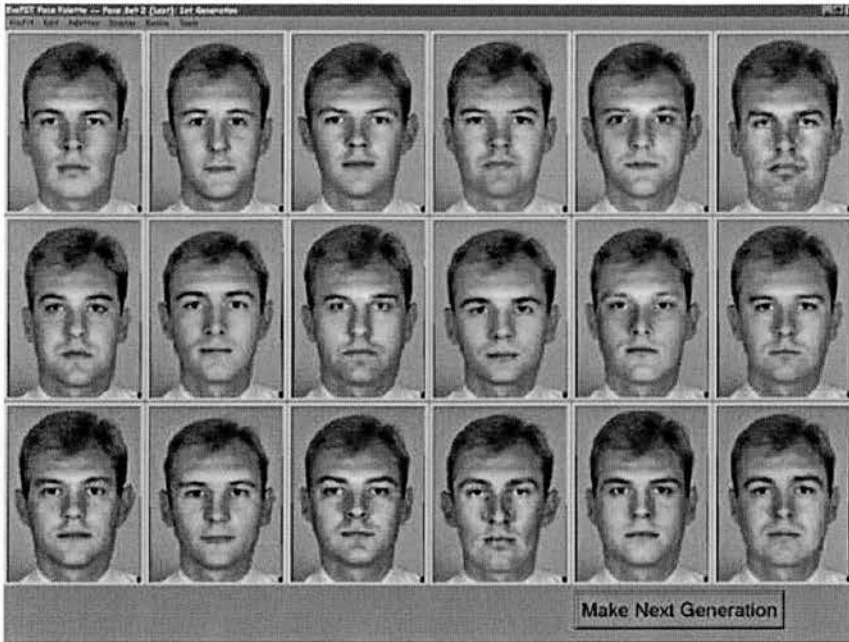


Figure 1.48 An example set of faces during face evolution, using the EvoFIT system for evolving faces for witness identification (From Frowd, 2000).

more similar to the target face. These methods have been incorporated into a computer program called EvoFIT (Figure 1.48), used to construct photo-fit images to aid criminal investigation (Frowd, 2000).

### **3D face manipulation**

Objects can be scanned to give a 3D representation, for example by using a Cyberscanner (Figure 1.49). Scans produced by this type of scanner<sup>18</sup> produce a depth map and a colour map (Figure 1.50).

Depth and colour maps can be manipulated (averaged, caricatured and transformed) in an analogous way to 2D images (Rowland, Perrett, Burt, Lee, & Akamatsu, 1997). For example, Figure 1.51 presents European and Japanese 3D head models with different levels of masculine characteristics, determined from averages of 3D male and female heads.



**Figure 1.49** A Cyberscanner in position to scan a person's face. The scanner head moves round the individual reading colour and depth information separately. (Image ©Cyberware Corp.)

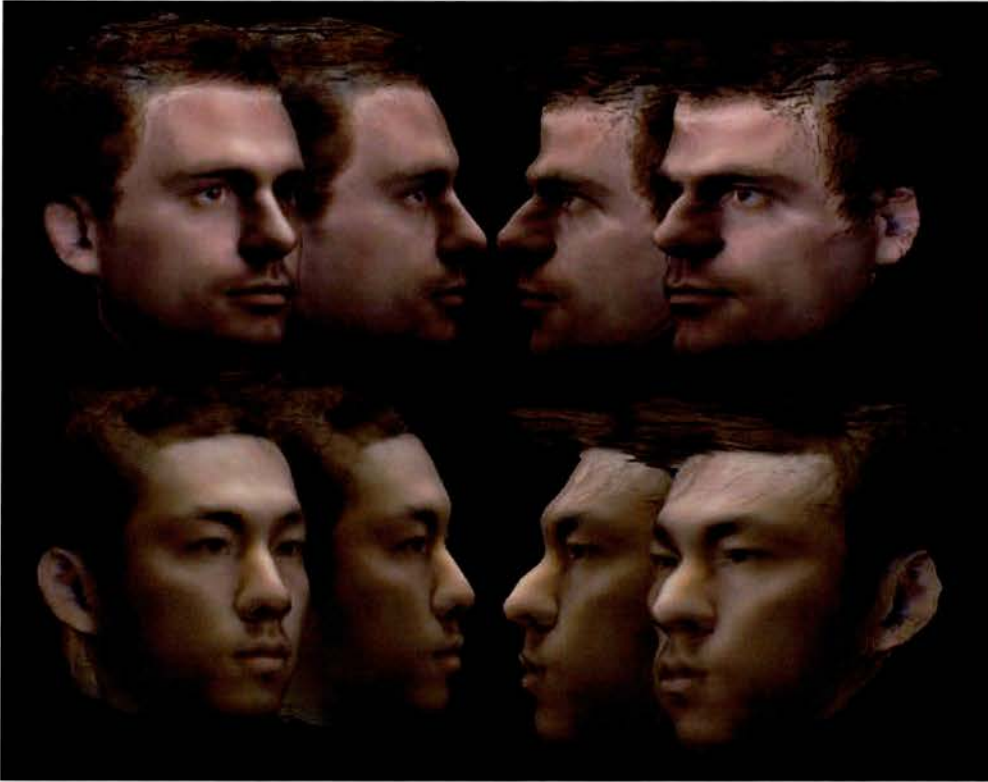


**Figure 1.50** The depth map (left; higher pixel intensity relates to smaller distance from sanner) and colour map (centre) produced by a Cyberscanner are reconstructed to give a view of the 3D representation (right). As the head of the scanner moves round the individual who is being scanned, it records depth and colouration. These can be expressed as two 'flattened cylinder images' or rewrapped round a virtual cylinder to produce a 3D image. (Image from Russell, 1997)

---

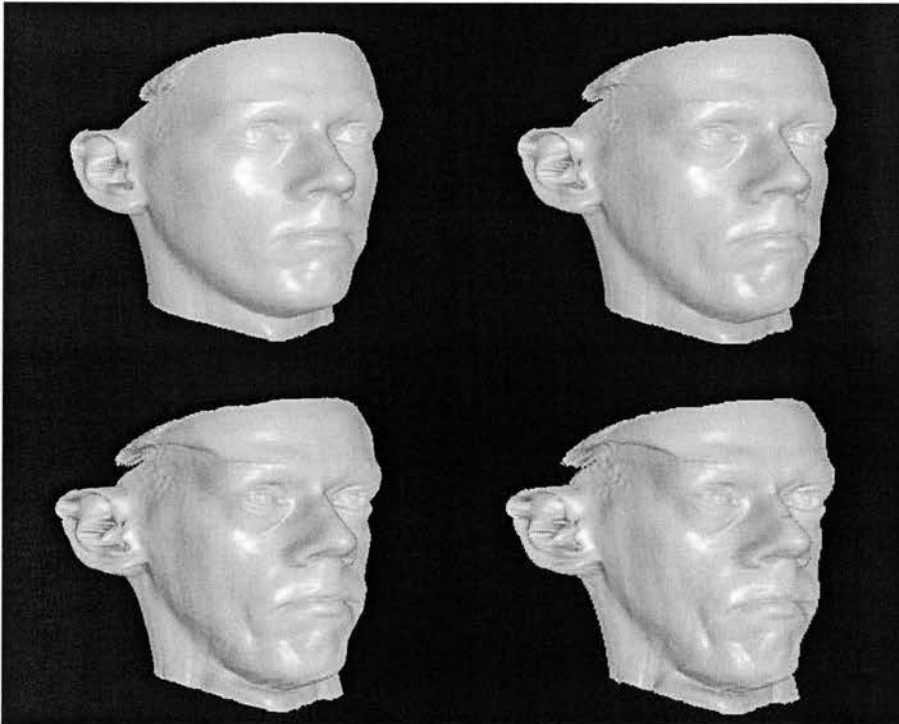
<sup>18</sup> For a review of other available means of collecting 3D scans, see Zhang (2001).





**Figure 1.51 Caricaturing masculinity in 3D shape. Above are presented 3D head views from feminised (left) and masculinised (right) 3D heads made from Caucasian (top) and Japanese (bottom) individuals. (Figure as used for Nature cover, Perrett et al., 1998)**

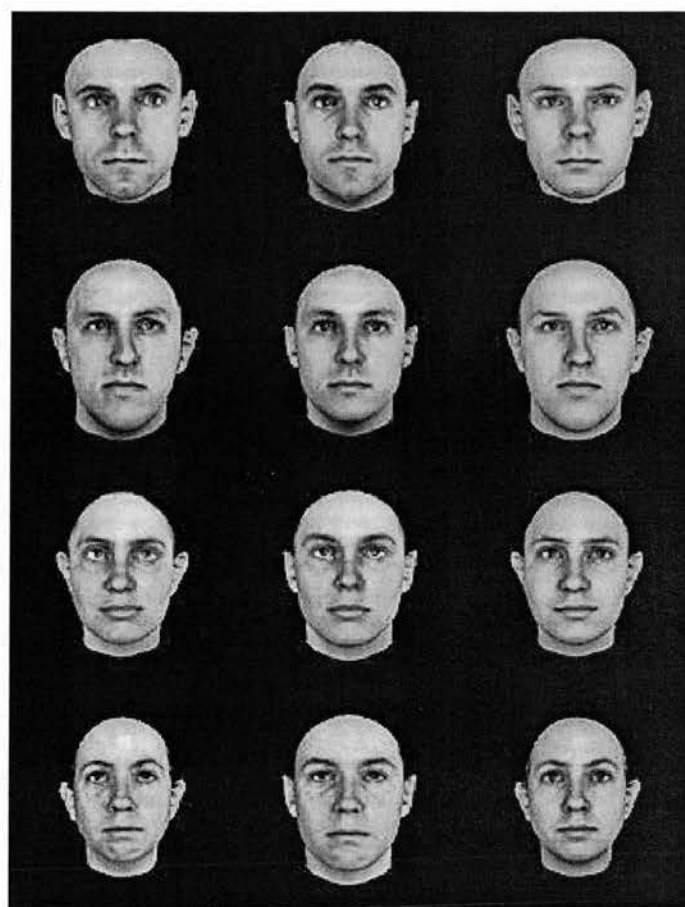
Vetter *et al.* have performed a number of recent psychology experiments on 3D human faces. Their methods include the use of ‘optical flow’, for image matching to find corresponding areas of the face (Vetter et al., 1997), PCA and the calculation of an approximate 3D model of the face from a 2D image (Blanz & Vetter, 1999; Vetter, 1998); or from a small number of feature points (Hwang, Blanz, Vetter, & Lee, 2000). These methods using 3D face data and PCA have also been applied to study face ageing, attractiveness and averageness (Deffenbacher et al., 1998; O’Toole, Price, Vetter, Barrlett, & Blanz, 1999a; O’Toole et al., 1997b), gender classification (O’Toole et al., 1997a), representation of faces across view (O’Toole, Vetter, & Blanz, 1999b); and to the investigation of models of face space (Blanz, O’Toole, Vetter, & Wild, 2000; Leopold et al., 2001).



**Figure 1.52**  
Caricaturing of identity in 3D PCA space has its most marked effect on apparent age. The originally scanned 3D face (top left) with 3 increasing levels of caricature (top right, bottom left and bottom right). (Images © O'Tool [www.utdallas.edu/~otoole/](http://www.utdallas.edu/~otoole/))

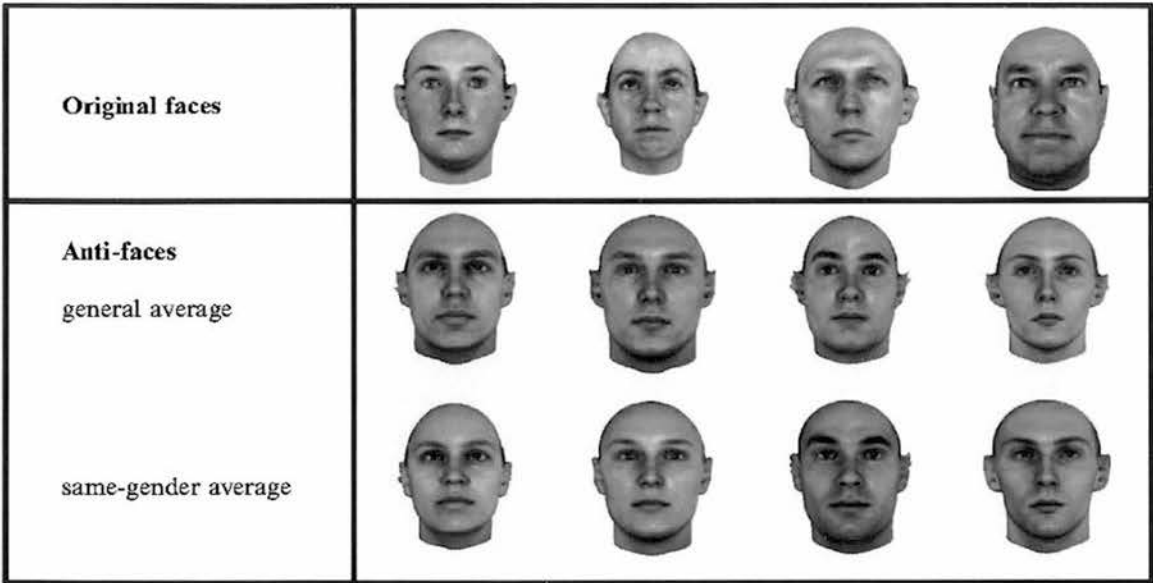
O'Toole *et al.* (1997b) caricatured 3D face shapes in PCA face space (Figure 1.52). They found that their caricaturing had the effect of increasing the perceived age of the face and suggest 3D caricaturing as a method of ageing faces. It is notable, from Figure 1.52, that the caricatured faces do not show many of the normal signs of ageing, like sagging of soft tissue. Distortions can make faces look older, possibly because older faces are more distinctive not only because of the lines and wrinkles that are not a feature of younger faces, but also from the differing effects of the environment (e.g. on some cartilaginous soft tissues like the nose and ear shape that continue to grow and whose shape becomes more gross and atypical with age). This does not mean that making a face more distinctive will necessarily make it age appropriately or predict how the individual might look in a few years time.

O'Toole *et al.* (1999b) in a later paper investigated the findings of Langlois *et al.* (1990; 1994) using 3D heads. Subjects rated the attractiveness of 3D representations of unaltered 3D heads, 3D heads with the average shape but the individual's original colour and 3D heads with the average colour but the individual's 3D shape (Figure 1.53). It was found that subjects rated both the average colour and the average shaped heads as being more attractive but also, in the



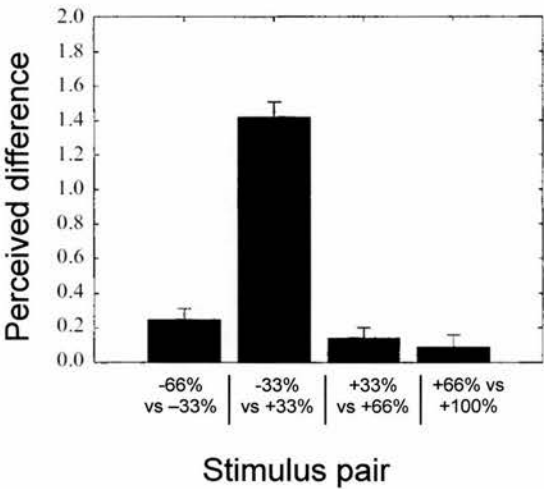
**Figure 1.53** Examples for face stimuli. Unaltered 3D faces (left column), with average 3D shape but original colour (centre column) and average colour information but original 3D shape (right column) (Images © O'Toole [www.utdallas.edu/~otoole/](http://www.utdallas.edu/~otoole/))

second experiment, as being younger. O'Toole *et al.* interpreted their findings as reflecting a preference for average faces since differences in perceived age did not account for the changes in attractiveness ratings.



**Figure 1.54** Position of original faces (top row) on each of the components returned from a principle component analysis was reversed to create anti-faces. This was performed when the analysis included both genders (middle row) and for when the analysis only included same gender faces (bottom row). (Figure from Blanz *et al.* 2000)

3D models have also been used to investigate face space. Recently, Blanz *et al.* (2000) extended the caricature method to create anti-faces in PCA face space. An anti-face lies at the same distance from the mean as the original face but in the opposite direction along each of the PCA components. Examples of the original faces



**Figure 1.55** Perceived difference between image pairs based on the distance from average.

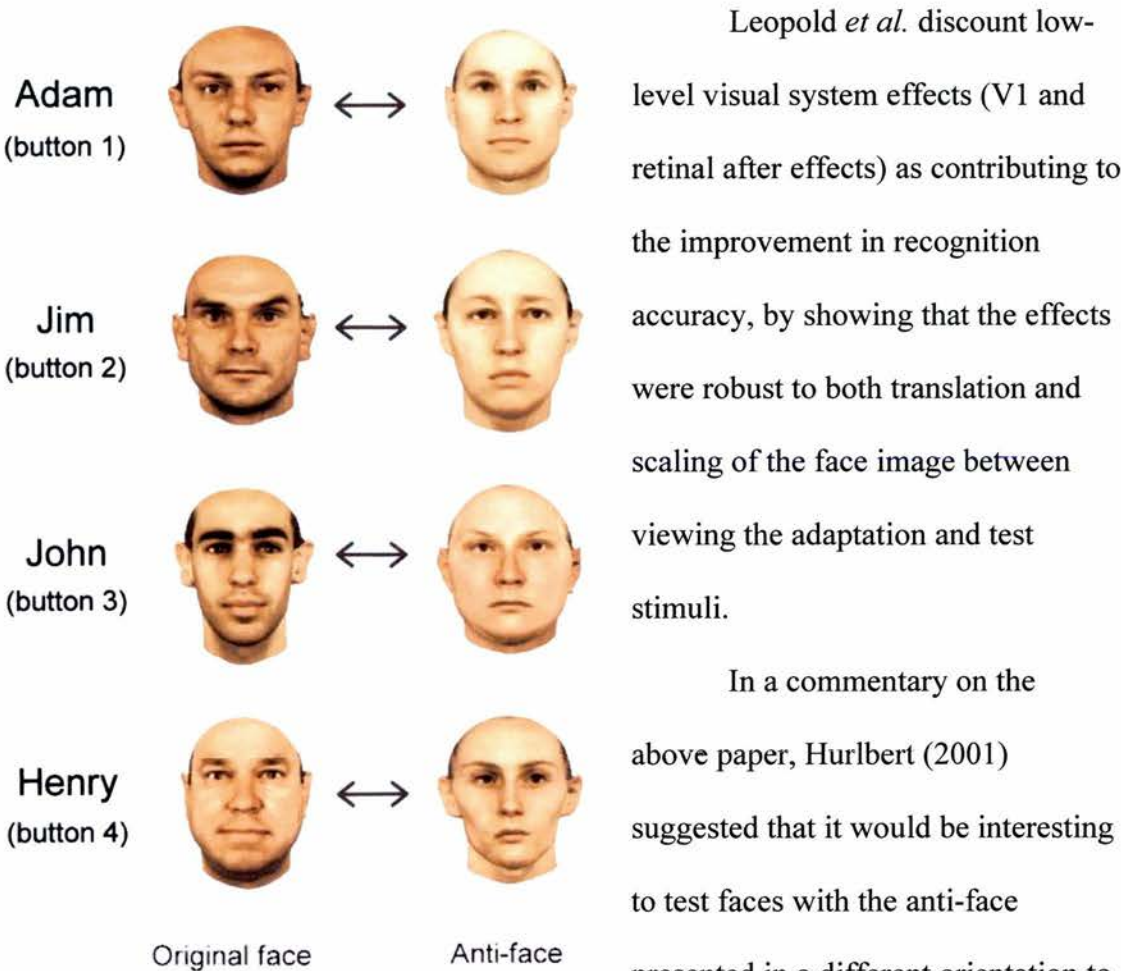
and anti-faces are shown in Figure 1.54. Subjects were asked to rate how different pairs of faces were. The pairs of faces came from a morph between original face images (+100%) , the average face (0%) and the anti-face (-100%). The full-blown anti-faces were not used because of glitches that tended to spoil them. Face pairs containing an

anti-face (-33%) and a face (+33%) were rated as being significantly more dissimilar than would be expected, if perception of differences were linear with respect to the amount of caricature (see Figure 1.55). Blanz *et al.* suggest that this represents a perceptual discontinuity between faces and anti-faces. This jump in dissimilarity could, instead of being a discontinuity, be related to the space being non-linear. Alternatively, at a neural level, more neurons may be responsive to faces that are near to average, since average faces may well encompass all of the most commonly found traits. Better discrimination of faces, given their linear difference in face shape, might therefore be expected for faces that are near to average, in a similar way that pairs of stimuli near a categorical boundary are better discriminated than equally different stimuli that are further from the categorical boundary (see section on image morphing). Indeed, the average face is probably, in many different ways, very near the categorical boundary. For example, it might be expected that the average face would have a nose size that would neither be classified as being big or small and might thus represent the categorical boundary between big and small noses. This would mean, for example, that a 33% face derived from an individual with a large nose, would have a nose that would be classified as large rather than small and the related 66% face would also have a nose classified in the same way as being large but the -33% face's nose would be described as small, as it is smaller than average for the particular group.

Using similar stimuli to Blanz *et al.* (portrayed in Figure 1.56), Leopold *et al.* (2001) investigated the influence of after-effects on the perception of face stimuli derived from four different identities. Subjects were trained over a period of weeks to recognise the four 3D faces until they were able to recognise the individual faces reliably when presented for short intervals (200ms). Subjects were then presented



with the anti-face, followed immediately by one of the 4 individuals that they had been trained to recognise. Following presentation of anti-faces, subjects were more sensitive to the presentation of the corresponding face but less sensitive in the recognition of the other individuals. The heightened sensitivity meant that correct recognition of the identity of an individual was more likely, even when faces had been made less distinctive by morphing the original faces towards average.



**Figure 1.56 Faces and anti-faces (and from Leopold et al., 2001)**

orientation. Leopold *et al.* did test for changes in orientation between the anti-face and the test face but found that subjects were distracted by the apparent rotation of the

stimuli. This suggested that it might not be possible to test the effect between inverted and non-inverted faces.

The authors interpret their results as being due to effects of viewing on higher order shape representations of faces, conveyed by neurons in an area known as the superior temporal sulcus (STS) becoming adapted to the anti-faces and therefore more sensitive to the corresponding target faces. Alternatively, the effect may be due to apparent movement focusing attention on significant features. During the experiment, the test stimulus face image follows directly after the adaptation face stimulus. This might produce a feeling of movement that might add salience to the cues that related to an individual face's identity. Indeed, when later in the paper the authors display the anti-face and the face with a gap between them, a blank, the size of the recognition advantage is diminished.

Thus it is possible that the motion of contours between corresponding anti-faces and faces of the same identity was responsible for the enhancement of recognition. Such a low or mid level visual recognition explanation is investigated in Chapter 7 using anti-expression stimuli from Chapter 6. These anti-expression stimuli are based on the 6 basic emotions and in Chapter 6 were used for the investigation of differences in perception of what is a neutral face. Chapter 7 uses these anti-expression stimuli to investigate a possible facilitatory effect of anti-expressions on the identification of subsequently presented expression faces. If such a facilitatory effect can be found, it would suggest that processes underlying the after effects described by Leopold *et al.* were more pervasive, applying to different facial descriptions.

## **Overview**

This chapter has presented an introduction to the computer graphics techniques used in contemporary face perception research. A short overview of the experimental chapters that form the rest of the thesis will now follow. Each of the experimental chapters starts with a brief review of the relevant psychological literature.

In the first experimental chapter, Chapter 2 of this thesis, the computer graphics technique of morphing is adapted to mix between pairs of face images to produce single stimuli for the investigation of possible brain processing, laterality effects. In the next chapter, Chapter 3, shape warping and blending techniques, combined with other computer graphic methods, are used to probe the finding in the facial literature that humans prefer less symmetric faces while non-human animals prefer more symmetric features. This chapter on symmetry and attractiveness then investigates possible mediators of preferences for symmetry; including an individual's own attractiveness and an index of their self rated well being.

Previous studies have shown that it is possible to capture cues relating to the perception of different facial attributes using blending. These cues can then be manipulated in other face images the by application of the information captured in the blend images. The utility of image warping, caricaturing and transforming techniques for the capture and manipulation of facial cues to skin health is investigated in the Chapter 4.

In this thesis single images of individual faces are studied and are a rich source of cues, but such images lack the dynamic information that is a normal part of face-to-face human interaction outside the laboratory. Previous studies (Berry, 1990; Kamachi et al., 2001; Yoshikawa & Wataru, 2001) have suggested that dynamic

actions influence the manner in which faces are perceived. Chapter 5 presents a foray into dynamic facial action and explores whether the combination of dynamic and static cues present in video footage can provide observers with information about whether an individual is enjoying consuming a food, the strength of the food's flavour, and which of the four basic tastes (salt, sweet, sour and bitter) predominates in the food. The cues that are communicated between a taster and an observer are discussed in terms of their utility for guiding food selection in humans and other species.

In the next chapter, Chapter 6, computer graphic based techniques are applied to devise a new test for the investigation of facial expression perception. The test is applied to investigate differences between control individuals and those with depression. Tests of facial expression perception employed in previous investigations of depression have make use of categorisation involving word labels, however internal emotional state is known to affect the choice of vocabulary used by subjects (Niedenthal & Halberstadt, 1997). So, previously reported effects may therefore reflect differences between groups in word salience, rather than differences in perception of emotionally expressive faces. Chapter 6 presents a test in which subjects are asked to make a face look as emotionally neutral as possible. Subjects' concept of what is neutral may be important since, if this is shifted, the interpretation of the whole of emotional face space may be different. Beck suggested that erroneous interpretation of ambiguous situations is of importance in depression (Beck, Rush, Shaw, & Emery, 1979). Neutral faces are potentially ambiguous as they necessarily do not portray the emotion that their wearer is experiencing.

In the penultimate chapter of this thesis, Chapter 7, anti-emotion stimuli from Chapter 6 are used to investigate a possible, short-term, perceptual after effect that

may be beneficial to expression recognition. Leopold *et al.* (2001) found that the viewing of the anti-faces facilitated the recognition of subsequently presented faces of individuals. Chapter 7 presents an investigation of a possible analogous effect for the recognition of facial expressions.

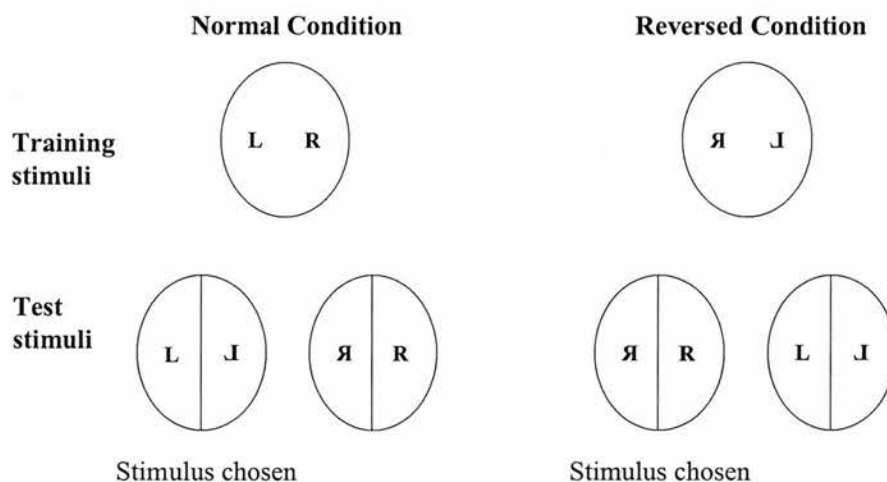
The face portrays many different types of information and diverse techniques are of use in the psychological investigation of these varied facial attributes. Thus the work presented in this thesis has employed a range of computer graphic techniques as reviewed in this introduction.



## **Chapter 2: Perceptual asymmetries in judgements of facial attractiveness, age, gender, speech and expression.**

## Introduction

This chapter explores the role of preferential attention paid to the left or right side of face stimuli when subjects are asked to make judgements relating to age, attractiveness, gender, lip-reading and facial expression. Historically this work started with the observation that has long been a part of folklore, that the side of the face to the viewer's left looks more like the owner than the side to the viewer's right. This observation was first scientifically investigated by Wolff (1933) who observed that chimaeric face stimuli produced by combining the left and mirror left of peoples' faces were thought by subjects to look more like the original people than chimaeric faces produced by combining the right and mirror right (Figure 2.1, Normal Condition). This bias was first thought to arise as a property of the face owner rather



**Figure 2.1: Training stimuli.** Ovals represent facial stimuli, L and R denoting the sides of the face to the left and right for the viewer, as originally photographed. Test stimuli: split ovals represent chimaeric faces made from one half of the training face and its mirror reflection joined at the vertical midline. Wolff (1933) found that subjects trained on the Normal Condition thought that the chimaeric stimulus made from the side of the face falling to the left ( L ) looked more like the training face than the chimaeric made from the side of the face falling to the right ( R ). In the Reversed Condition subjects were trained with a mirror image of the whole training face. Note that reversed L and reversed R denote the sides of the training stimulus in relation the original photographic image. Gilbert and Bakan (1973) showed that the bias in matching identity was due to perceptual asymmetries in the viewer because subjects trained on the Reversed Condition still chose the chimaera made from the half face which was on their left during training (denoted by a reversed R).

than a property of the viewer. It was speculated that the left side of a person's face was more 'public' in character; i.e. more representative of the individual. Gilbert and Bakan (1973) dispelled this speculation by training subjects to name unfamiliar faces, some of which had been mirror reversed (Figure 2.1, Reversed Condition). They found that judgements of likeness of chimaeric stimuli to training stimuli depended on which side of the face was on the observer's left side during training. Thus Gilbert and Bakan found the bias in facial likeness was due to asymmetries in the perception of faces rather than arising from physical asymmetries expressed by individual faces. Several authors define the sides of a person's face with respect to the face owner (Gilbert & Bakan, 1973; Wolff, 1933; Zaidel, Chen, & German, 1995) but, since the perceptual bias described above is due to asymmetries in the viewer, the sides of face-stimuli will be referred to from the viewer's perspective (except where stated explicitly otherwise).

Visual perceptual asymmetries may be due to the predominance of the right hemisphere in facial identification since it is more intimately connected to primary visual areas responsible for the processing of visual information from the left hemifield (Gilbert & Bakan, 1973). Experiments using chimaeric stimuli have also shown that the observer's right hemisphere is not only more influential in processing facial identity than the left but also predominates in other areas of facial processing including perception of facial expression (Christman & Hackworth, 1993; Rhodes, 1993; Schiff & Truchon, 1993) and gender (Luh, Redl, & Levy, 1994).

Why the right hemisphere predominates in the processing of facial identity, expression and facial gender is unclear. Neurological evidence implicates the right hemisphere in the processing of faces. For example, most, if not all, prosopagnosic patients have sustained damage to their right hemisphere (De Renzi, Garthia, &

## Production Asymmetries

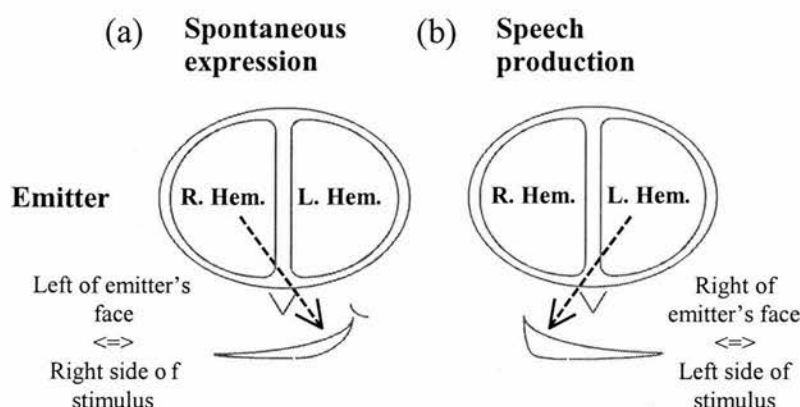


Figure 2.2: The main direction of information flow in emitter. Both spontaneous expression and speech production are asymmetrical. This is caused in (a) by the stronger motor outputs to the contralateral side of the face from the 'more expressive' right hemisphere (R. Hem.) and in (b) by the stronger outputs to the contralateral side of the face from the 'more linguistic' left hemisphere's (L. Hem.).

Faglioni, 1989). The selective attention to the left of faces found when testing facial chimaeras may be a reflection of the right hemisphere being more efficient than the left hemisphere at the processing of facial material *per se*. Alternatively, the right hemisphere may be better at the analysis of the spatial configuration of any visual pattern and, because of this, predominates in face processing (Rhodes, 1993). A different explanation is that asymmetrical patterns of eye scanning from left to right (arising from reading habits) may result in the left hemifield receiving more attention than the right hemifield (for review see Rhodes, 1986). This scanning bias should result in preferential attention to the left hand side of all visual stimuli.

Judgements on facial dimensions could also be biased by inspection of the side of the owner's face which normally portrays more physical information relevant to a given judgement. Facial movements made both during spontaneous expressions and during talking are asymmetrical in the amount of information portrayed (Figure 2.2a, b). Facial expressions are more intense on the lower left half of emitter's faces (Borod, Caron, & Koff, 1981) and the right side of the brain has been argued to play a greater part in the control of emotional expression (Borod et al., 1981). Indeed, it may be through the stronger contralateral motor outputs that the right hemisphere predominates in influencing the lower left half of the owner's face (see Figure 2.2a).

Since posed smiles tend to be symmetric and spontaneous smiles asymmetric, the left side of the owner's face will also be informative as to whether a smile is genuine (Wylie & Goodale, 1988). So from a viewer's perspective the right half of face images should be more important in the judgement of real emotions. Previous studies (Christman & Hackworth, 1993; Rhodes, 1993; Schiff & Truchon, 1993) have, however, found that perception of facial expression is biased towards information occurring on the left side of face images.

Facial movements during talking are biased to the right side of the owner's face (Figure 2.2b). Most people (76%) tend to talk with a greater amplitude of mouth movements on the right side of their mouth than on their left. This is presumably due to the greater involvement of the speaker's left hemisphere in language production and the stronger connection between left hemisphere motor control systems and the musculature of the left side of the speaker's face (Graves, Goodglass, & Landis, 1982; Wolf & Goodale, 1987). As the speaker's right falls to the left of the viewer, the left side of the face images should be more informative about speech sounds when talking and it may be expected that viewers would attend more to the left side of stimuli when lip-reading.

This idea is supported to some extent by the results of one previous paper investigating the laterality of lip-reading. In two experiments, Campbell (1986) demonstrated a right hemisphere advantage (left visual hemifield) for lip-reading. In both experiments tachistoscopic presentation was used (200 ms stimulus duration in the first experiment and 100 ms in the second). The right hemisphere bias found in this experiment, could, however, have been due to other factors. Sergent has claimed that the use of fast presentations with visually complex stimuli, that are hard to discriminate, favours right hemisphere visual processing strategies (Sergent, 1982).

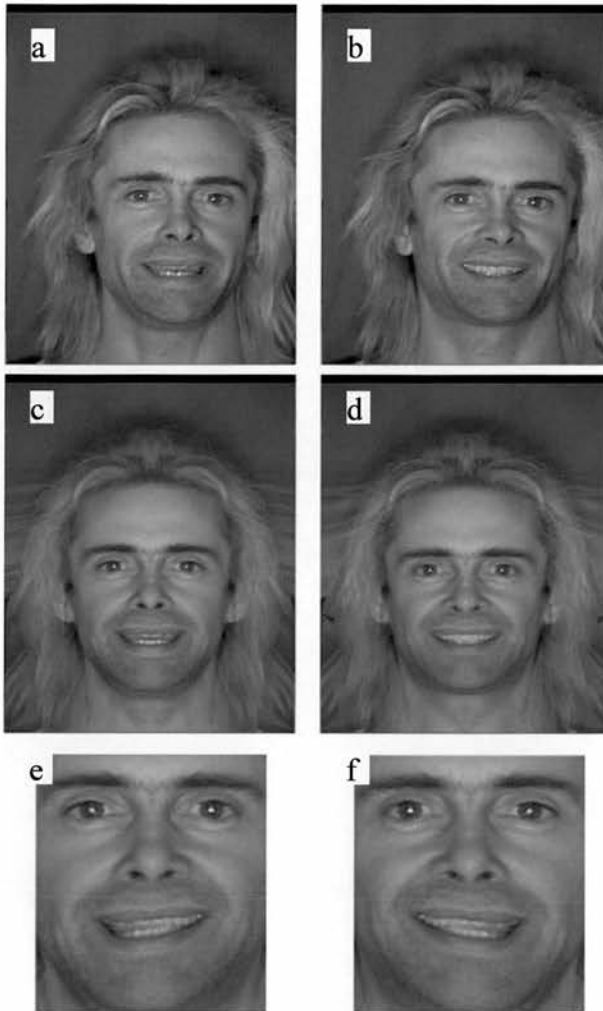


The relationship between presentation time and hemispheric bias is, however, complicated and other researchers have suggested the opposite (see Nicholls, 1995).

Neuropsychological evidence suggests a critical role for the left hemisphere in the perceptual aspects of lip-reading. Campbell *et al.* (1987) investigated two individuals with lateralized posterior brain lesions. Mrs D had right occipito-temporal lesions and severe face recognition problems (prosopagnosia). The other individual, Mrs T, had left occipito-temporal lesions and severe reading deficits (alexia). Campbell *et al.* found that the right hemisphere lesioned patient was able to lip-read normally but that the left hemisphere lesioned patient was unable to lip-read. This evidence, combined with the finding that fast presentation of stimuli may itself cause right hemisphere bias in normal viewers (Sergent, 1982), raises the possibility that the use of chimaeric stimuli under free vision (i.e. prolonged exposure) to examine lip-reading, could lead to a left hemisphere advantage and a right hemistimulus perceptual bias.

This chapter sets out to investigate the relative importance of the left and right sides of face stimuli in the processing of a variety of facial dimensions. The experiment looks at facial expression and gender, and three additional facial dimensions: age, attractiveness and lip-reading. It was expected that the bias in attending to information present on the left side of facial stimuli, found previously for expression and gender, would be replicated and would also be manifest in judgements about attractiveness and age.

Lip-reading stimuli were included that required observers to match speech sounds to the shape of the mouth in single video frames taken mid-utterance. Since such judgements are more linguistic or verbal in nature, these judgements may depend on aspects of processing for which the left hemisphere is dominant. It was predicted,



**Figure 2.3: The manufacture of stimuli.** Each stimulus was made from two face images. In the case of the first lip-reading these were (a) an individual saying 'ee' and (b) the same individual saying 'i'. These face images were first made symmetrical (c and d) before being merged together and cropped to form the final composite images. The first of these images (e) has on the left a symmetrical face saying 'ee' and on the right the symmetrical face saying 'i'. The second (f) is the mirror image of the first.

therefore, that the left hemiface perceptual bias present for other facial judgements would disappear or would be reversed for lip-reading stimuli.

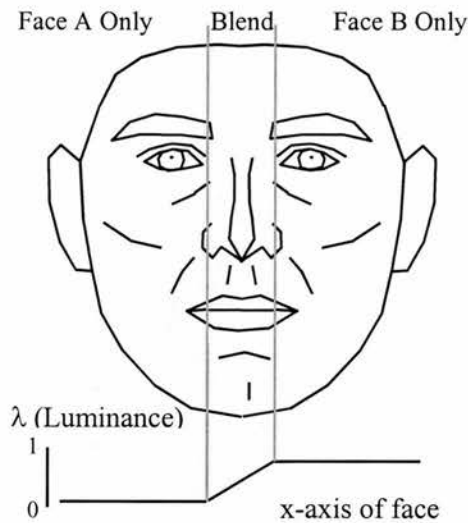
## **Method**

### **Subjects**

The stimuli were tested on 132 subjects of whom 24 reported being left handed. 73 of the subjects were female and 59 male with an average age of 20.7 years.

### **Manufacture of stimuli**

Two face images were used to manufacture each pair of test stimuli. These face images (Figure 2.3a, b) were first made symmetrical (Figure 2.3c, d). This



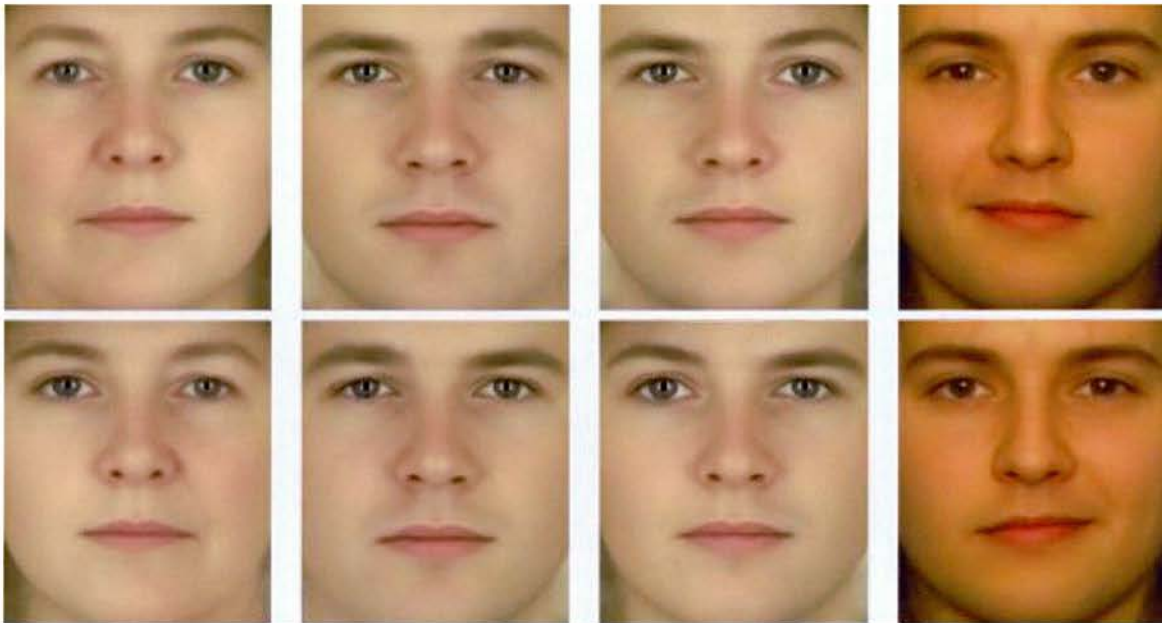
**Figure 2.4:** The mask used to manufacture chimaeric faces. The luminance ( $\lambda$ ) of the mask varied from black (0) on the left to white (1) on the right with a ramp across the midline. This is illustrated as a graph at the base of the figure. The position of the mask is defined relative to the average face shape. As can be seen the ramp was matched in width to the average nose. Chimaeric stimuli were made by taking shape and colour information from 2 faces in proportion to the value of 1 (see text).

preliminary procedure removed all structural asymmetries present in the face. Then, using information contained in the 'mask image' (Figure 2.4), the two symmetrical face images were merged together with a gradual change in shape and colour from one face image to the other across the vertical mid-line, producing the first stimulus of the pair (Figure 2.3e). The second stimulus of the pair was then made by reflection of the first stimulus across the y-axis (Figure 2.3f). The stimuli thus produced were pairs of chimaeric faces where the left and right halves differed in a particular dimension.

## Description of stimuli

Stimuli were manufactured for 5 facial dimensions: ageing (2 stimuli: *male ageing* and *female ageing*), attractiveness (4 stimuli: *male employee*, *female employee*, *male model* and *female model*), gender (2 stimuli: *gender employee* and *gender model*), expression (2 stimuli: *smile-neutral* and *sad-neutral*) and lip-reading (2 stimuli: *ee-i* and *ee-ss*). In total, this gave 12 stimulus types. Examples of some of these stimuli are presented in Figure 2.5.

The first ageing stimulus, *male ageing*, was made from two male face blends; one was composed of 21 faces, average age 22 years, and the other of 18 faces,



**Figure 2.5: Test stimuli. Four examples of these are presented. Top row of stimuli, from left to right: age (*female ageing*), the left half is a blend of old female faces and the right is a blend of young female faces; attractiveness (*male employee*), the left half is a blend of faces rated high in attractiveness and right is a blend faces rated low in attractiveness; gender (*gender employee*), the left half is a blend of male faces and the right a blend of female faces; emotion (*smile-neutral*), the left half is a blend of smiling faces and the right is a blend of the same faces with neutral expression. The bottom row of stimuli are mirror images of the top row stimuli. Testing involved pairs of corresponding stimuli from top and bottom rows.**

average age 51 years (for further details of these blends see Burt & Perrett, 1995). The second ageing stimulus, *female ageing*, was made from two female blends. The first was composed of 34 faces, average age 22 years, and the second of 17 faces, average age 52 years.

The first of the attractiveness stimuli, *male employee*, was made from two male face blends, a ‘*high*’ attractiveness blend and a ‘*low*’ attractiveness blend. The *high* was composed of the 15 most attractive faces from a population of 59 males aged 20-30 and the *low* from the 15 least attractive faces. The original faces were presented in random order and rated for attractiveness on a 7 point Likert scale (1 very unattractive, 7 very attractive) by Caucasian subjects (15 male, 15 female, aged 18-30) (Perrett, May, & Yoshikawa, 1994). So that all of the features of the two blends coincided across the midline, the *low* face blend was stretched in height enabling the

distance between the mouth and the midpoint between the eyes to match those of the *high* blend. The second attractiveness stimulus, *female employee*, was made from two female face blends, a *high* blend (15 most attractive female faces from a population of 60, aged 20-30) and a *low* blend (15 least attractive faces from the same population) matched in height to the *high*. The original Likert ratings of female faces were obtained from Caucasian subjects (26 female, aged 18-45 and 10 male, aged 19-24) (Perrett et al., 1994). The shape (Perrett et al., 1994) and the colour (Burt and Perrett, unpublished work, see Chapter 2 Appendix) of the *high* male and female blends were perceived by most subjects as more attractive than the *low* blends. The third and fourth attractiveness stimuli, *male model* and *female model*, were made from different collection of faces (see Rowland & Perrett, 1995). The *male model* was made from two male face blends, a *high* blend (11 faces rated to be most attractive from a population of 43 faces taken from fashion magazines) and a *low* blend (11 least attractive from the same population), stretched in height to match the *high*. The *female model* was composed of two blends of female faces, a *high* blend (20 faces rated as most attractive from a population of 61 taken from fashion magazines) and a *low* blend (20 faces rated least attractive from the same population), stretched in height to match the *high*. Likert scale ratings of the original male and female fashion model faces were made by 40 Caucasian subjects (35 female 5 male aged 19-30).

The first gender stimulus, *gender employee*, was made from the *high* attractive male face blend used to make the *male employee* and the *high* attractive female face blend used to make the *female employee*. The female face blend was stretched so that the vertical distance from the centre of the mouth to between the eyes was the same as the male blend. The second gender stimulus, *gender model*, was made from the *high* attractive male face blend used to make *male model* and the *high* attractive female



blend used in *female model*. Again the female face blend was stretched so that the vertical distance from the centre of the mouth to the midpoint between the eyes was the same as the male blend.

The first expression stimulus, *smile-neutral*, was made from a blend of 6 smiling faces and the same 6 faces with neutral expressions. The second expression stimulus, the *sad-neutral*, was made from two grey-scale blends, one of 8 sad faces and the other the same 8 faces with neutral expressions (original sad and neutral faces were taken from Ekman & Friesen, 1976).

The first lip-reading stimulus, *ee-i*, was made from two single frames of a face. One frame was taken while the subject was pronouncing the word 'bee', captured mid utterance during the 'ee' sound. The second frame came from the same face pronouncing the word 'side', captured during the 'i' sound. The second lip-reading stimulus, *ee-ss*, was made from single frames from sequences of the face saying 'bee' and 'hiss' captured during the sounds 'ee' and 'ss'. Pilot work showed that the faces from which the lip-reading stimuli were made were easily understood by subjects, who recognised around 90% of them.

## **Manufacture of symmetrical face images**

The co-ordinate positions of 208 'feature points' (Benson & Perrett, 1991), each representing a specific feature of the face (e.g. nose, right of lip), were logged. This spatial description of the face was then transformed so that the eyes were in standard positions, equidistant from the vertical midline. The symmetrical face shape was then calculated by averaging the co-ordinate position of each 'feature point' with the co-ordinate position of the related feature point on the opposite side of the face reflected across the y-axis. The starting face image was then warped into the new

symmetrical shape. The image was then blended with its mirror image to produce a face symmetrical in shape and texture.

### **Manufacture of left-right chimaeric face images**

After being made symmetrical, the two faces were merged together across the mid-line in shape and in RGB colour. The mask (Figure 2.4) encodes which areas of the face are to be taken from the face image 'a' (dark areas on the mask) and which from the face image 'b' (light areas). The two faces were blended together (colour and shape) in varying proportions across the midline according to the intensity of the mask.

Calculation of the final face shape proceeded in a stepwise fashion from the first to the last feature point. From the mask, the intensity of the pixel at each feature point was looked up (using the feature point data for the average face). This defined the fraction of fade ( $\lambda$ ) for the feature point (i.e. how much of the information for that point was taken from the first vs. the second face). This provided the spatial data for a new chimaeric face shape with the features on the left side defined by one face (e.g. old) and the right side defined by a second face (e.g. young). The two face images were warped into this chimaeric shape and then integrated in a pixelwise fashion. Starting with the first pixel the fraction of fade ( $\lambda$ ) was again looked up from the mask and used to calculate how much of the RGB colour information for that pixel in the final image would come from the first (old) vs. the second (young) warped face images.

Alternatively the method can be looked upon as an equation where  $\lambda$  is the luminance (0-1 in 256 steps) derived from the mask image using the associated shape data, new is the new (x,y) co-ordinate or RGB triple (for calculation of shape and

colour respectively) and a and b are the (x,y) co-ordinate or RGB triple for the first and second face images respectively:

$$\text{new}(x,y / \text{RGB}) = a(x,y / \text{RGB}) - \lambda (b(x,y / \text{RGB}) - a(x,y / \text{RGB}))$$

The final stimulus was made by cropping the output image to a size of 268 by 308 pixels. The second stimulus of the pair was made by reflection of the first. Then each facial stimulus was printed (size: 4.6 cm by 5.2 cm, 150 dpi 24 bit colour, Printer: Mitsubishi Sublimation S3410-30, CMYK 4 colour print roll).

## Procedure

During testing the pair of stimuli were placed one above the other. Subjects were asked to choose whether the lower or higher stimulus was older, happier, more feminine, more attractive, or more like it was saying 'ee', 'i' or 'ss', depending upon the dimension that the stimulus tested. Each subject was tested with all of the stimuli in a self paced manner. The vertical position of each stimulus within each pair was counterbalanced between subjects. For all facial dimensions, the adjective used to test subjects was also counterbalanced (i.e. for age judgements, half the subjects were asked which is older and half were asked which is younger).

## Results

For most dimensions tested, the left hand side of the faces biased the subjects' perceptual judgements. That is, information on the left of the stimuli had a dominant influence on responses. This bias was significant for perception of age (*male ageing*, 77% of subjects' judgements were based on the left of stimulus faces, Binomial test,  $p < 0.05$ ; *female ageing*, 73%,  $p < 0.05$ ). The bias to the left side of stimuli was

significant for attractiveness judgements for the first two facial stimuli (*male employee*, 67%,  $p < 0.05$ ; *female employee*, 60%,  $p < 0.05$ ) but did not reach significance for the second two attractiveness stimuli (*male model*, 57%,  $p = 0.07$ ; *female model*, 55%,  $p = 0.13$ ). Significant biases to the left of stimuli were found for judgements about gender (*gender employee*, 66%,  $p < 0.05$ ; *gender model*, 69%,  $p < 0.05$ ) and expression (*smile-neutral*, 58%,  $p < 0.05$ ; *sad-neutral*, 59%,  $p < 0.05$ ).

The perceptual bias was reversed for the choice of lip-reading stimuli. Subjects showed a significant tendency to make choices dominated by information present on the right hand side of the face stimulus (*ee-i*, 42%,  $p < 0.05$ ; *ee-ss*, 42%,  $p < 0.05$ ) rather than the left side.

The current study was not designed to investigate the role of handedness; nonetheless, self-reported handedness was recorded. Of the subjects, 18% reported being left handed. The effect of handedness (estimated in this way) on lateral bias in facial judgements was analysed in a 2-way ANOVA with stimulus type (12 levels) and subject handedness (2 levels) as main factors. Analysis again revealed a significant main effect of stimulus type ( $F_{1,11} = 5.41$ ,  $p < 0.05$ ) but there was no significant difference between lateral biases for right and left-handed subjects ( $F_{1,11} = 2.11$ ,  $p = 0.17$ ) and no interaction.

## **Discussion**

Forced choice, free vision testing of 132 subjects confirmed and extended the view that, for many aspects of face processing (perception of age, attractiveness, gender and expression), subjects preferentially attend to information on the left side of the facial stimulus. In the case of lip-reading the opposite bias was found; most subjects' perception of which sound was being spoken was influenced by the information on the right side of the face (Figure 2.6).

## Perceptual asymmetries

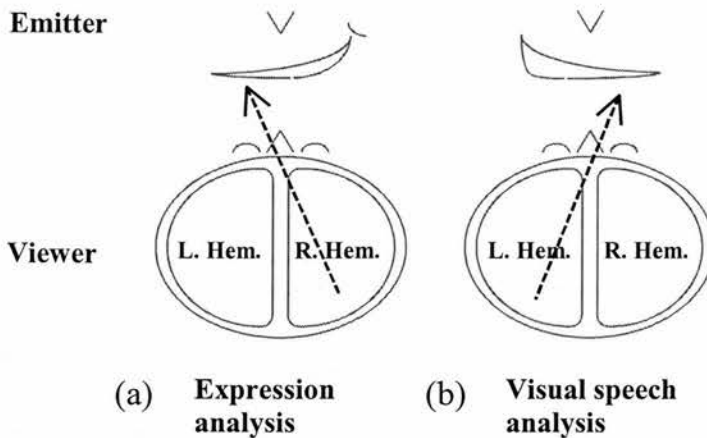


Figure 2.6 The main direction of information flow in observer. Rather than taking note of the side of the face which contains the most information (see Figure 2.2), perception of expression (a) is biased to material on the right of the emitter and visual perception of speech (b) is biased to information on the left side of the emitter. This may be caused in (b) by the dominant role of the viewer's right hemisphere in face perception and the stronger inputs from the left side of the image to this hemisphere and in (b) by the dominant role of the viewer's left hemisphere in visual speech analysis and the stronger inputs from the right side of the image to this hemisphere.

## **Expression, gender and age**

As predicted, the experiment replicated the findings of perceptual bias to the left half of the chimaeric face stimuli under free vision for processing of both expression (Christman & Hackworth, 1993; Rhodes, 1993; Schiff & Truchon, 1993) and gender (Luh et al., 1994). These perceptual biases were interpreted as reflecting the predominance of the right hemisphere in the processing of facial information.

Of more interest was the finding of a perceptual bias to the left half of the face stimulus during judgements of age. A right hemisphere predominance in perception of facial age might be anticipated from the neurological evidence of De Renzi *et al.* (1989) who found that patients with right posterior brain injury performed worse at face age tests than patients with other brain injury.

## **Attractiveness**

Since attractiveness is strongly influenced by age and the degree of femininity or masculinity of a face, it may not be surprising that judgements of attractiveness were also biased to the left side of stimulus faces. It should be noted, however, that though a significant bias was obtained for two stimuli (*male employee* and *female*

*employee*), the bias did not reach significance for two other stimuli (*male model* and *female model*). This may have been due to the difference in attractiveness between *high* and *low* attractiveness groups being smaller for the *models* than the *employees*. Alternatively, it could be due to subjects making preference judgements on the basis of the entire hemistimulus picture colour rather than on the basis of hemiface. Colour was tightly controlled in the conditions under which the images for the *employee* stimuli were photographed so colour differences between the left and right of the *employee* stimuli were therefore minimal. However for the *model* stimuli control over the photographic conditions was not as tight leading to noticeable differences in overall colour cast across the left and right hand side of the *model* stimuli.

Zaidel *et al.* (1995) found that chimaeric faces made by combining the right side of the owner's face and mirror right (Figure 2.1, Left composite from the viewer's perspective) were found more attractive than chimaeric face composites made from the left side of the owner's face and mirror left (Figure 2.1, Right composite from the viewer's perspective). Chimaeric composites made from the owner's left (viewer's right) were judged to have more pronounced "smiling expressions" than composites made from the owner's right (viewer's left). This led Zaidel *et al.* (1995) to speculate that: "attractiveness and smiling may be asymmetrically and oppositely organised on the face".

The results also indicate that perception of attractiveness is lateralized and, in accordance with the interpretation of Zaidel *et al.*, that there is a perceptual bias favouring the left side of the face stimuli (right side from the owner's perspective). From the results, the perception of expressions, including smiling, also showed a bias to information on the left side of face images. This finding would appear to contradict the interpretation by Zaidel *et al.*, but replicates previous findings of a left sided bias



(right hemisphere advantage) in the perceptual processing of expressions under free fixation (Christman & Hackworth, 1993; Rhodes, 1993; Schiff & Truchon, 1993).

Symmetric smiles can be perceived as less genuine when compared with asymmetric expressions (Wylie & Goodale, 1988) and look less attractive (Kowner, 1996). Thus chimaeric stimuli may have unnatural looking symmetric expressions even when constructed from natural faces. The greater attractiveness of left face chimaeric stimuli in the experiment of Zaidel *et al.* may therefore be due to the fact that these stimuli were perceived as having less pronounced symmetrical smiles (Kowner, 1995).

Whatever the interpretation of the findings of Zaidel *et al.*, the experiment indicates that for judgements about attractiveness there is a perceptual bias with greater attention being paid to information on the left side of both male and female face stimuli. Aesthetic judgements about faces appear to exhibit the same lateral bias as judgements about face gender, expression and age.

## **Lip-reading**

Neurological evidence (Campbell, 1987) suggests that subjects use their left hemisphere during the recognition of speech sounds from facial images (Figure 2.6b). In the experiment, normal subjects were allowed free fixation of stimuli. Thus, a strong inference cannot be made that information from one side of the facial stimuli projects primarily to the contralateral cerebral hemisphere. Nevertheless the bias to attend to information on the right side of stimuli may be interpreted as reflecting a predominance of the left hemisphere's abilities during processing. In this sense the current findings are consistent with the neurological evidence of a predominant role of the left hemisphere in processing facial cues to speech sounds. The use of long presentation times (unlimited exposures) may have uncovered a left hemisphere bias,

which had been overshadowed in previous work with normal subjects (Campbell, 1986) by fast presentation times, which Sargent (1982) has argued favours the right hemisphere (though see Nicholls, 1995). Recent work has shown that lateralization of linguistic processing depends both upon the specific task demands and the individual (see Hickok & Poeppel, 2000, for discussion). Although some studies show results that suggest a right hemisphere bias for linguistic tasks most show a left hemisphere bias.

Recently there have been attempts to explain cognitive psychological phenomena, including language, in terms of a direct link between perception and actions (e.g. see Preston & de Waal, 2000). Explanations of this type have gained ground with the finding of 'mirror neurons' in Area F5 of the macaque monkey (Gallese, Fadiga, Fogassi, & Rizzolatti, 1996). Mirror neurons are neurons that fire to both the individuals making a motor action and the sight of the action being performed by another individual. The macaque area F5 is a likely analogue of Broca's area in humans and humans show evidence of mirror neurons (e.g. Rizzolatti et al., 1996). These neurons are located in the frontal hemispheres with connection to both motor and sensory areas. Theorists have suggested that mirror neurons are of pivotal importance in the evolution of the extension of imitative gestures, that is language (Rizzolatti & Arbib, 1998). It is possible that 'mirror neurons' in the left hemisphere may be used when performing the 'lip-reading' task and that these neurons have greater input from primary visual areas on the left side of the brain. This preferential access to the primary visual areas of the left hemisphere could create the right side bias observed here for lip reading stimuli.

An alternative hypothesis, that humans learn to attend to the side of the face in which more information about the sound is displayed, is disproved by the results,

since the right side of the face (left of stimulus, Figure 2.2b) shows a greater amplitude of movement when talking (Wolf & Goodale, 1987), but is more informative for interpreting speech (Wolf & Goodale, 1987). So learning to pay attention to the most informative side does not have a large effect if any, on the perceptual processes underlying lip-reading. If lip-reading used the most informative side of the face then a left side of face advantage would be expected, which is the opposite of the result obtained here.

### **Hemispheric asymmetries in processing**

Milner and Dunne (1977) presented chimaeric faces briefly (100 ms) and found that normal subjects behaved somewhat like commissurotomed ('split-brain') patients when the junction of the two half faces at the vertical mid-line was obscured from view with a black strip. With such stimuli, many subjects were unaware that the two sides of the chimaeras were different in identity. They experienced perceptual completion of a single intact face. When subjects were required to match the identity of the chimaeric faces by pointing with their left hand, identity matching was significantly biased to the left half faces. When a verbal response was required in a different block of trials, the subjects showed a non-significant bias to name the right half faces (i.e. projecting to the left hemisphere). Thus hemispheric biases in face processing were influenced by the mode of response (verbal label or pointing with the left hand).

A critical aspect of Milner and Dunne's (1977) experiment was the presence of the vertical strip at the mid-line to prevent subjects seeing the join between the two half faces. In the experiment, the employment of blended stimuli and graphics processing allowed the two half faces to be gradually merged in the chimaeras. This again meant that the mid-line junction was not obvious. Indeed subjects were unaware

that the two sides of the face were different in the dimension being tested (i.e. they did not realise one side was young and the other was old). Furthermore most subjects did not realise that the two stimuli, presented as a pair, were mirror images. If subjects realise that each chimaera is a composite of two separate faces, then they may not engage the same type of processing as that employed if the stimulus is seen as one image. This difference in processing of face identity was noticed with chimaeric stimuli made of two half faces of famous individuals joined at a horizontal mid-line (Young, Hellawell, & Hay, 1987).

The left sided perceptual bias with chimaeric faces is normally attributed to a right hemisphere processing superiority. It is not clear, however, why the right hemisphere bias arises. It is commonly attributed to a right hemisphere specialization in face processing. It could be, however, that all visual stimuli requiring complex configurational judgements produce right hemisphere bias. Alternatively, right parietal mechanisms involved in control of spatial attention may be selectively engaged during visual processing; these could cause a bias to scan the left side of all visual stimuli. From both of these latter explanations, the perceptual bias to the left side of faces need not reflect a right hemisphere specialization for face processing. Indeed, Luh *et al.* (1994) reported a right hemisphere advantage in perceptual judgements for both chimaeric face stimuli and chimaeric non-face patterns.

The lip-reading condition demonstrates that it is possible to obtain a perceptual bias to the right as well as the left side of face stimuli. To my knowledge such a right-sided bias has not been observed before in studies using free fixation. If these perceptual biases reflect hemispheric biases then evidence for specialization of the left and right cerebral hemispheres can be obtained with simple testing methods. The hemisphere predominantly engaged in processing is determined by the facial

dimension to be judged. For normal subjects the left hemisphere tends to predominate during processing of facial information about speech (lip-reading) and the right hemisphere seems to predominate during processing of other facial dimensions (age, gender, expression and attractiveness). Both of these findings are concordant with the neuropsychological studies of brain-damaged subjects.

## **Chapter 3: Facial attractiveness and symmetry.**



## ***Introduction***

In this chapter a series of experiments are presented. Experiments 1 and 2 were aimed at examining why, contrary to the literature relating to other animals, studies of humans suggested that manipulations to increase symmetry of faces resulted in stimuli that were less attractive. Experiments 3 and 4 move on to examine the mechanisms that bring about symmetry preferences and individual differences that relate to the strength of preference for symmetry. This chapter begins with a brief summary of the literature as it stood before the completion of Experiments 1 and 2.

In the last decade, the idea that symmetry confers fitness advantages in sexual competition across the animal kingdom has stimulated some exciting research in both humans and other animals. Asymmetries within an organism, particularly fluctuating asymmetries (e.g. Thornhill and Gangestad 1993)<sup>19</sup>, are widely thought to be the result of developmental instability resulting from disease, malnutrition, other environmental stresses encountered by the organism, and genetic factors effecting growth (for review see Møller, 1997). Symmetry has therefore been proposed to be an indicator of high developmental stability and history of health (Gangestad, Thornhill, & Yeo, 1994; Thornhill & Gangestad, 1993; Thornhill & Gangestad, 1999). Theories of sexual selection suggest that individuals are best off, from an evolutionary point of view, if they select the healthiest mate (Trivers, 1972).

The presence in the real world of the link between developmental stability and symmetry is supported by many correlational studies (review in Møller, 1997) that have shown greater asymmetries are caused in non-human animals by environmental

---

<sup>19</sup> Fluctuating asymmetries are symmetries that occur at random in a population such that on average direction of the asymmetry in the population is zero.

stresses (caused by habitat disturbance Lens & Van Dongen, 2000; Parsons, 1992) and genetic endowment (Møller & Thornhill, 1997). Symmetry appears to be weakly heritable (for review see Leung & Forbes, 1996; Møller & Thornhill, 1997) and symmetrical males are preferred mates in many species (for review see Moller & Thornhill, 1998).

It is possible that in choosing more symmetrical mates individual animals, like those in the above studies, are using cues other than symmetry since other cues to health may correlate with symmetry. Studies have therefore been performed in which the symmetry of individuals has been artificially manipulated and the effects on the animals' attractiveness as a mate recorded. The majority of these symmetry manipulation studies suggest that animals prefer symmetrical mates (see Moller & Thornhill, 1998). One rather clever study to uncover the role of symmetry in mate selection was performed by Swaddle and Cuthill in 1995. In this study, two possible signals, plumage loss and symmetry were manipulated. The chest plumage of male zebra finches, *Taeniopygia guttata*, was removed in either a symmetrical or asymmetrical pattern. Female zebra finches were found to prefer males who had more symmetrical plumage even if they had had more plumage removed than males with asymmetrical plumage. Thus symmetry may, under certain circumstances, be preferred over other indicators of condition.

Symmetry has also been studied in humans. Humans, like other animals, also show a link between developmental stability and symmetry (reviewed in Thornhill & Moller, 1997). For example, mentally retarded individuals show greater asymmetries than individuals without mental retardation (Malina & Buschang, 1984), and individuals who have been infected with parasites tend to be more asymmetrical than uninfected individuals (Bailit, Workman, Niswander, & Maclean, 1970).

Correlational studies have also shown that in humans, as other animals, more symmetrical individuals are preferred (e.g. Grammer & Thornhill, 1994). The greater attractiveness of symmetrical human individuals is not just confined to the lab. Preferences for more symmetrical human individuals appears to translate into a real world difference between less and more symmetrical individuals, as more symmetrical individuals report having more sexual partners, and having been younger at the time of their first sexual experience (Gangestad, Bennett, & Thornhill, 2001; Thornhill, 1993). Partners of more symmetrical males report having more orgasms, a finding that is explained by the authors in terms of the female orgasm being a mechanism for the preferential retention of sperm (Thornhill, Gangestad, & Comer, 1995). Females possessing more symmetrical breasts are both more sought after and more fecund (Møller, Soler, & Thornhill, 1995). Additionally, more symmetrical individuals score higher on measures of 'cultural intelligence' than less symmetrical individuals (Furlow, Armijo-Prewitt, Gangestad, & Thornhill, 1997).

In humans, as in other animals, the correlation between symmetry and other attractiveness cues means that symmetrical individuals might be preferred because of other attractiveness cues rather than symmetry<sup>20</sup>. To summarise, in non-human animals, and in humans there is a literature that suggests that developmental instability causes asymmetry. In non-human animals, the literature suggests that symmetrical individuals are preferred. Studies of non-human animals show that generally, in experiments that manipulate symmetry, more symmetrical individuals are preferred. At this point though, the non-human animal and human literatures

---

<sup>20</sup>It is noteworthy that the degree that symmetry is used as a cue to attractiveness is currently under contention in the human literature (Kowner, 2001 ; Penton-Voak et al., 2001; Scheib, Gangestad, & Thornhill, 1999).

diverge. Until the completion of Experiments 1 and 2, the literature on human attractiveness suggested that manipulations to increase symmetry in humans actually had the effect of lowering facial attractiveness (Kowner, 1996, 1997; Langlois, Roggman, & Musselman, 1994; Swaddle & Cuthill, 1995). The reason for this surprising disparity between the non-human and human literature related to mate selection might be found in the techniques used in the experiments to manipulate symmetry in humans or in aspects of the human face that are preferred in their asymmetrical form.



**Figure 3.1 Normal and symmetrical faces. The original face image (top left) can be distorted so that the positioning of its features is symmetrical (top right; using Experiment 2 methods). An alternative means of making the face symmetrical is to mirror down the image midline to produce a double left (bottom left) and double right (bottom right). Distortion of the nose is clearly visible on the images on the bottom row.**

In the aforementioned human symmetry manipulation studies, all but one manipulated symmetry using a ‘mirroring’ technique. An example of this technique is illustrated in Figure 3.1, and is similar to the technique used by Wolff (1933) in the previous chapter. While the technique does result in a symmetrical image, this

symmetrical image is a distorted representation of the original. This is exemplified in Figure 3.1. In this set of mirrored images, the nose has become much wider than it was in the original (top left) in one symmetrical image (bottom left) and much narrower in the other (bottom right). It may be argued that this is because the position of the mirror line is not bisecting the nose correctly, but the correct bisection line for one feature is not always the correct bisection line for the other features in the face. Such mirroring techniques have raised other criticisms e.g. skin moles or freckles and even hairstyles look unnatural or abnormal after this mirroring technique has been applied. Kownner (1996; 1997) attempted to reduce these problems by using computer graphic techniques to touch-up these abnormalities. However, even symmetrical face stimuli that have been touched up in this way were, with the exception of old male faces, found less attractive than the original less symmetrical images (Kownner, 1996; Kownner, 1997).

Swaddle *et al.* (1995) generated symmetrical images by using a morphing technique (see Chapter 1, section entitled “Image blending using feature points”) to blend original images with their mirror image. In this technique the original image is morphed with a mirror reversed version of itself so that half way through the morph the resulting image is symmetrical, being 50% of the original image and 50% of the mirror reversed image. Unlike the technique above, this technique results in midline features that are representative of the original face image. However, distortions will occur if the face image is rotated in the image plane; see Figure 3.2. Other oddities that diminish the naturalness of the final image are the duplication of features like spots so that they occur on both sides of the face, but with half the intensity of the original image. More serious distortions occur, however, if the landmarks that are used to transform the face are not at corresponding positions on either side of the face



image. A gross example of this is illustrated in the bottom line of Figure 3.2. Using this new computer graphics technique, Swaddle *et al.* (1995) found that making face images more symmetrical decreased their attractiveness.



**Figure 3.2.** Possible distortions induced when using morphing to make images symmetrical. To demonstrate possible distortions induced when using morphing to make images symmetrical, the original image (top) was rotated in the image plane (upper-left) before being averaged with a mirror image version of itself to produce a symmetrical version (upper-right). A distortion, compression of the face along the y-axis and expansion in x-axis can be observed. If points are laid incorrectly on the face image, even within the line of the feature, this can also cause a distortion. To demonstrate this the points representing the chin and corners of the lips are misplaced (bottom left). When the image is then blended with its mirror image, distortion of both the lips and chin result (bottom right).

As indicated above, a number of technical difficulties may have hampered studies exploring the relationship between manipulated symmetry and attractiveness in humans. However, putting these issues aside, there are possibly other explanations why individuals might prefer less than perfectly symmetrical faces. There may be certain aspects of the human face that are preferred in their asymmetrical form. Unlike



some parts of the body, the human face has no directional asymmetries in its resting structure (see Experiment 2 Preliminary Discussion for more details). This lack of directional asymmetries is not always true of the human face when it is not at rest. These asymmetries are evident in the face during speech: 76% of people tend to make movements of greater amplitude on the right side of their mouth (Graves, Goodglass, & Landis, 1982; Graves, Strauss, & Wada, 1990; Wolf & Goodale, 1987).

Asymmetries also appear to be present during emotional expressions. Overall studies have found that the left side of the face is more emotionally expressive than the right (Skinner & Mullen, 1991). There is debate about whether spontaneous expressions are more or less symmetrical than posed expressions (Skinner & Mullen, 1991) though, most studies (Borod, Caron, & Koff, 1981; Borod & Koff, 1983; Wylie & Goodale, 1988) have found that spontaneous expressions are more asymmetrical than posed expressions. The seeming spontaneity of asymmetrical expressions could possibly be a reason why symmetrical face images might not seem as attractive as the original asymmetrical face images.

To test whether retaining these small, expressive asymmetries, while making the rest of the face symmetrical would result in a more preferred face shape, stimuli were manufactured that were symmetrical apart from their oral region. Thus, in the experiment below, subjects chose between original (asymmetrical) face shapes, wholly symmetrical face shapes and face shapes that were symmetrical in all features but their mouth. These face shapes were taken from the images of individuals but, in order to avoid the problems commented on above related to features like spots, each face shape was provided with the average symmetrical texture from a group of face images.

## ***Experiment 1 blended texture symmetry and mouth manipulation***

### **Aims**

Previous experiments suggest that symmetrical face shapes are not the most attractive to humans. This may be because symmetrical mouths are unattractive since spontaneous or felt expressions tend to be asymmetrical and therefore preferred or because of inappropriate methods used to make the face stimuli symmetrical. In order to test this, preferences for original face shapes (*'normal'*), symmetrical face shapes (*'symmetrical'*) and symmetrical face shapes with original mouth shapes (*'mouth'*) were compared. It is hypothesised that if the symmetrical mouths are unattractive then the *mouth* stimuli will be preferred over the *symmetrical* stimuli. If symmetrical features are preferred then the *symmetrical* stimuli will be preferred over the *normal* and *mouth* stimuli.

### **Methods**

#### **Subjects**

11 male (average age 30) and 11 female (average age 27) completed the experiment.

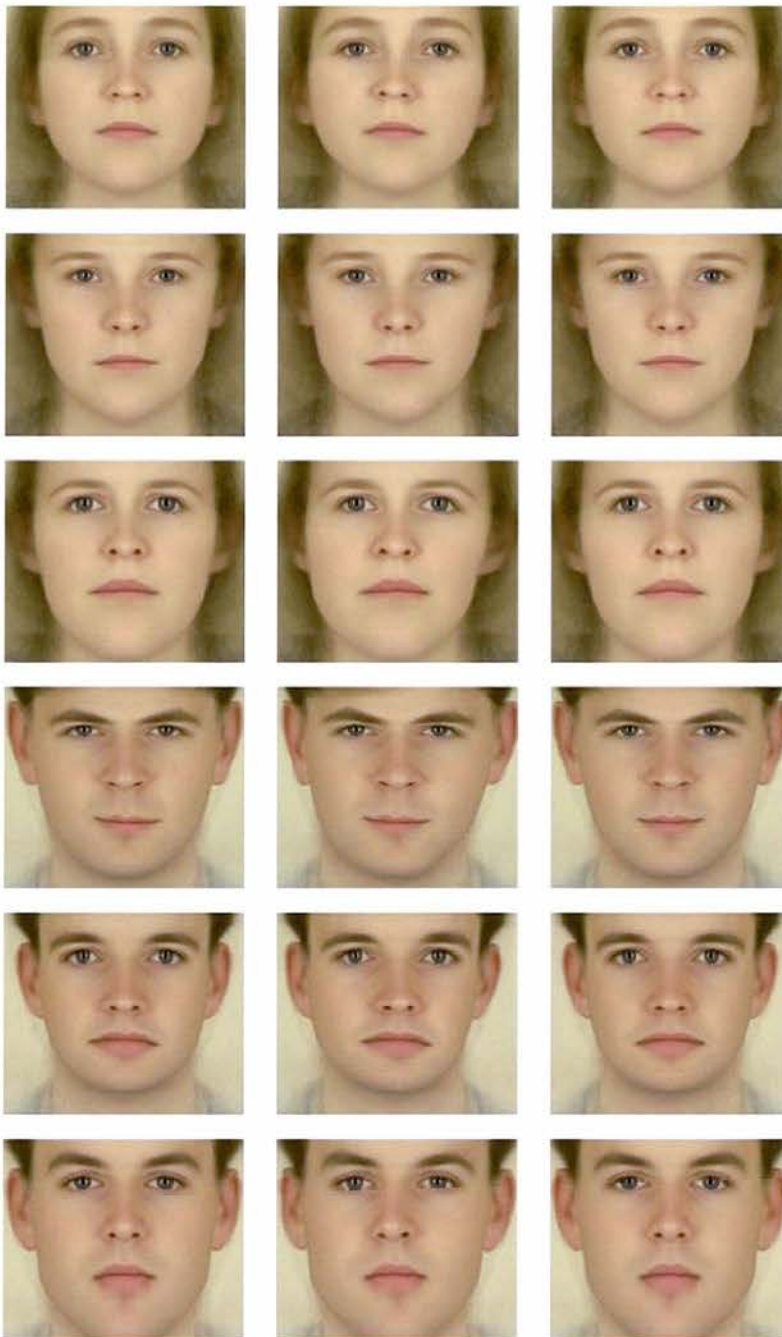
#### **Stimuli**

The face shape of a set of adult, 24-bit face images that had been photographed with a neutral pose facing directly toward the camera was delineated (Benson & Perrett, 1991). This face set has been previously used in other papers (Burt & Perrett, 1995; Perrett, May, & Yoshikawa, 1994). All of the individuals

photographed were free from make-up and other facial adornments, wore hair off the forehead, and males were clean-shaven. Using the delineation data, the face images were first normalised. Normalization involved the translation, scaling and rotation of the face to match the eyes to standard eye positions in which the positions of the centre of the eyes are horizontal and equidistant from the image midline.

To manufacture the texture for the stimuli, a subset of 20 male and 20 female face images were made symmetrical. The shape of each face was made symmetrical by replacing each feature point with the average of it and its corresponding point reflected in the midline. The original image was first warped into this new symmetrical shape to produce an image that was symmetrical in feature position. This image was then blended with a reflected version of itself to produce a symmetrical face image. These 20 male and 20 female symmetrical face images formed the basis of the texture of the experimental stimuli.

To make the stimuli themselves, 32 female and 32 male images were selected from the set of adult face images. *Normal* stimulus versions for each of the original face images were made by blending the 20 symmetrical face images of the same gender into the original (normalised) face shape. *Symmetrical* stimulus versions of each face image were manufactured by creating a symmetrical version of the normalized face shape, and then blending the 20 symmetrical face images of the same gender into the symmetrical face shape. A *mouth* version of the face was made by swapping the points related to the mouth shape from the *normal* face shape into the *symmetrical* face shape and then blending the 20 symmetrical face images of the same gender into the new shape. These images were then cropped and printed out onto cards. Thus were created 64 *normal*, 64 *mouth* and 64 *symmetrical* stimuli examples of which are presented in Figure 3.3.



**Figure 3.3** Examples of stimuli presented in Experiment 1. On the left are presented *normal* stimuli that preserve the original individual's face shape, on the right are the corresponding *symmetrical* stimuli and in the centre the *mouth* stimuli with symmetrical features except for their mouth shape.

## Presentation

Subjects were presented with 2 blocks of trials: male stimuli and female stimuli. The order of the blocks and trials within the blocks were randomised. At the start of each trial the 3 stimuli: *normal*, *symmetrical* and *mouth*, derived from one individual's shape, were laid out in random order in front of the subject and the

subject asked to select the stimulus that they found most attractive. This stimulus was taken away and the subject asked to indicate from the remaining stimuli the face image that they found most attractive. The order that the stimuli were selected was recorded.

## Results

Each trial was examined in terms of 3 comparisons: whether the subjects preferred the *normal* to the *symmetrical*, the *symmetrical* to the *mouth* or the *normal* to the *mouth*.

### Female stimuli

Both male and female subjects chose the *symmetrical* stimulus as being more attractive than the *normal* stimulus on more than half of the 32 trials (one sample t-test against an expected mean of 16; female subjects:  $av=24.7$ ,  $sd=4.8$ ,  $t_{10}=6.06$ ,  $p<0.001$ ; male subjects:  $av=21.55$ ,  $sd=5.8$ ,  $t_{10}=3.19$ ,  $p<0.01$ ). Subjects tended to prefer the *symmetrical* stimulus to the *mouth* stimulus, although this was only significant for female subjects (t-test; female subjects:  $av=20.46$ ,  $sd=6.0$ ,  $t_{10}=3.47$ ,  $p<0.05$ ; male subjects:  $av=18.5$ ,  $sd=4.0$ ,  $t_{10}=3.04$ ,  $p=0.68$ ). Neither male nor female subjects significantly preferred the *mouth* to the *normal* stimuli (t-test; female subjects:  $av=17.5$ ,  $sd=3.7$ ,  $t_{10}=1.33$ ,  $p=0.26$ ; male subjects:  $av=17.5$ ,  $sd=5.3$ ,  $t_{10}=0.90$ ,  $p=0.39$ ).

### Male stimuli

Both male and female subjects chose the *symmetrical* stimulus as being more attractive than the *normal* stimulus on more than half of the trials (one sample t-test against an expected mean of 16; female subjects:  $av=25.0$ ,  $sd=5.2$ ,  $t_{10}=5.72$ ,  $p<0.001$ ;

male subjects:  $av=21.6$ ,  $sd=4.8$ ,  $t_{10}=3.86$ ,  $p<0.01$ ). Subjects tended to prefer the *symmetrical* stimulus to the *mouth* stimulus although, this was again only significant for female subjects (t-test; female subjects:  $av=20.5$ ,  $sd=6.0$ ,  $t_{10}=3.45$ ,  $p<0.05$ ; male subjects:  $av=17.3$ ,  $sd=4.1$ ,  $t_{10}=1.41$ ,  $p=0.19$ ). Both male and female subjects preferred the *mouth* stimulus to the *normal* stimulus (t-test; female subjects:  $av=23.7$ ,  $sd=3.9$ ,  $t_{10}=5.72$ ,  $p<0.001$ ; male subjects:  $av=20.3$ ,  $sd=4.8$ ,  $t_{10}=3.30$ ,  $p<0.01$ ).

## Preliminary discussion

Results of Experiment 1 demonstrate that people do prefer the *symmetrical* faces to both the original, *normal* face shapes and partially asymmetrical *mouth* face shapes. The preference for the *symmetrical* stimuli over the *mouth* stimuli lead us to conclude that there is no support for the hypothesis that in resting faces asymmetrical emotional expression cues may make the asymmetrical faces preferable to symmetrical faces. This hypothesis will therefore not be investigated further as a possible reason for asymmetry being preferred in humans.

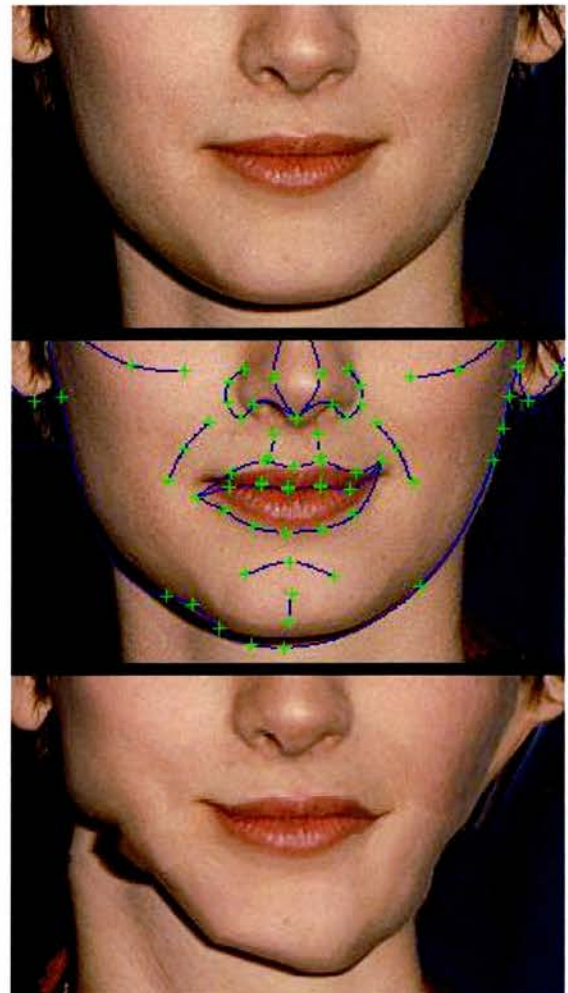
The results of Experiment 1 challenge the unlikely disparity between the non-human and human literature relating to mate selection and symmetry. The results suggest that the disparity is due to flaws in the techniques of previous experiments manipulating human facial symmetry (Kowner, 1996; Kowner, 1997; Langlois et al., 1994; Samuels et al., 1994; Swaddle & Cuthill, 1995). The introduction highlighted some of the possible problems with previously employed techniques. There are also criticisms that can be levelled at the techniques of Experiment 1.

Two possible problems arise with the stimuli used in Experiment 1. Firstly, although points were placed with care in the manufacture of the face shapes, any deviation from the proper feature point position would cause the *normal* face to look asymmetrical in a manner that was not representative of the original face shape. These



types of unnatural asymmetries may possibly drive the preference for symmetry in the results of Experiment 1. Secondly, some of the asymmetries in the experiment may have been caused by head rotation of the individual being photographed. These original photographs would have possessed information from shading that would have enabled individuals studying the original images to correct their perception of the facial asymmetries taking into account head rotation. The symmetrical textures that were mapped onto the face shapes on the other hand did not possess these cues.

Subjects judging these faces might interpret any head rotation as anatomical asymmetries in the face. To avoid these possible confounds, a second experiment was performed in which the original face image were compared with face images remapped to deform the facial features into a more symmetrical shape. This comparison is more conservative since the graphic changes are smaller, any 3D shape depth information from shadows will still be present in the manufactured image and slight misplacement of the feature points will result in an asymmetry in the image manufactured to be more symmetrical (Figure 3.4).



**Figure 3.4.** The methods used to make faces more symmetrical, will if the points are not placed correctly result in a less symmetrical image. A gross example is illustrated above. An original image (top) is delineated with points incorrectly placed (middle). When the shape of the face is made symmetrical and the texture warped into this shape a less symmetrical face results (bottom). The results of this method can be compared to the symmetrical face stimulus resulting from the same incorrect point placement as shown in Figure 3.2.

## ***Experiment 2 original texture***

### **Aims**

It was postulated that the results of Experiment 1 would be replicated by using a different, more conservative, technique. Accordingly, it is hypothesised that subjects will prefer the symmetrically shaped versions of the faces to the original versions of the faces.

### **Methods**

#### **Subjects**

49 subjects (12 male) age between 18 and 22.

#### **Stimuli**

From the same pool of face images used in Experiment 1, 13 female and 13 male face images were selected. Two versions of each face image were manufactured: a *normal* and a *symmetrical* version. *Normal* stimuli were the unaltered original images apart from being normalised. Normalizing the images involved first translating, scaling and rotating the face shape to standardize eye position before warping the original face image into the new shape. The standardized eye positions placed the eye centres horizontal on the image and equidistant from the image midline. The shape of the normal image was then made symmetrical and the *symmetrical* stimulus made by warping the original face image into this new shape. Both images were then cropped to give the final stimuli (illustrated in Figure 3.5).



**Figure 3.5: Examples of stimuli presented in Experiment 2. In the top row are presented the *normal* stimuli, the original images normalized on eye position. On the bottom row the *symmetrical* stimuli, the original images have had their feature positions remapped to be symmetrical in shape. In the symmetrical stimuli the ‘texture’ of the image remains as it was in the original so features like spots remain asymmetrical.**

## Presentation

Each pair of faces, *symmetrical* and *normal* were presented side by side (in random order and counterbalanced for side of presentation) to subjects who made a forced-choice comparison of attractiveness. The male and female face pairs were presented in separate block of trials. After performing the experiment, each subject was debriefed and asked what they thought the basis of their judgements had been.

## Results

Subject preferred the *symmetrical* stimulus to the *normal* stimulus in 57.8% of the trials. This was significantly above chance (*t*-test against a hypothesised mean of 50%,  $t_{48}=5.22$ ,  $p<0.0001$ ). An ANOVA was performed on the number of symmetrical faces chosen by each subject with *subject sex* as one factor and *face sex* as the other. There were no significant effects of *subject sex* ( $F_{1,47}=1.58$ ,  $p=0.215$ ) or *face sex* ( $F_{1,47}=1.380$ ,  $p=0.246$ ) and no significant interaction ( $F_{1,47}<0.0001$ ,  $p=0.987$ ).



At debriefing only 25% of subjects were aware that symmetry had been manipulated in the stimuli or had thought that it might have affected their judgements. In a further analysis with these subjects excluded, *symmetrical* faces were still preferred to *normal* face stimuli (*t*-test against a hypothesised mean of 50%,  $t_{29}=3.24$ ,  $p<0.005$ ).

## **Preliminary discussion**

The results of Experiment 2 confirm the findings of Experiment 1: more symmetrical face images are preferred. Thus, Experiment 2 used face images from a normal Western European population and a more conservative image manipulation technique (see Experiment 1, Preliminary Discussion) to show that subjects prefer symmetrical faces even when subjects appear unconscious of the nature of the symmetry manipulation.

Originally, when planning Experiment 1, it was thought that previous research had not found facial symmetry to be preferred in human faces because some asymmetries present in resting faces may be preferred to symmetry. Although, it seems that this is not the case it was decided to examine more thoroughly the asymmetries present in the population of faces used for the experiments presented here. Asymmetries are classified into three types: directional asymmetry, bipolar (or anti-symmetry) and fluctuating asymmetry. These are excellently illustrated and reviewed by Kowner (2001). Directional asymmetries, like the heart position, occur when the direction of the asymmetry is common to the population. Bipolar asymmetries occur in hand size, in which individuals tend to have either a larger right or left hand (Kowner, 2001). Neither directional asymmetry nor bipolar asymmetry is generally thought to indicate developmental instability. Fluctuating asymmetries, random levels of asymmetry with a mean of zero in the population, are thought however to relate to developmental instability.

Although, there is little previous evidence of visible directional asymmetries in the human face (but see, Burke, 1971; Ferrario, Sforza, Ciusa, Dellavia, & Tartaglia, 2001; Keles, Diyarbakirli, Tan, & Tan, 1997; Previc, 1991), it was decided to test for the presence of directional asymmetries in the faces used. Vertical and horizontal differences in position were analysed between left and mirror right feature positions in the eye position normalised faces (see Table 3.1). For 12 of the 13 feature point coordinates there was no significant directional asymmetry. The amount of fluctuating asymmetry (average of unsigned asymmetries) was always far greater than the amount of directional asymmetry (average of signed asymmetries). There is little evidence of directional asymmetries in the face images that were used and no directional asymmetries greater than 1 pixel, the distance between the eye centres being set at 170 pixels. The single significant directional asymmetry was in the height of the eye corners which showed a small (0.7 pixel) directional asymmetry. This apparent directional asymmetry is very small compared to the other asymmetries observed. It is therefore concluded that the results of the experiments are due to asymmetries that fluctuate.

Point discription	t=	p=	Fluctuating asymmetry	Directional asymmetry
outside corner of eye (x)	-1.98	0.057	1.0	-0.3
outside corner of eye (y)	3.81	0.001	1.2	0.7
inside corner of eye (x)	-1.38	0.178	1.3	-0.3
inside corner of eye (y)	1.37	0.182	0.9	0.2
center of nose (x)	0.25	0.804	5.0	0.2
eyebow corner out (x)	-0.11	0.916	1.7	0.0
eyebow corner out (y)	1.56	0.130	3.1	0.9
eyebow corner inner (x)	0.28	0.782	2.6	0.1
eyebow corner inner (y)	0.68	0.501	1.6	0.2
lip corner (x)	0.64	0.530	3.7	0.4
lip corner (y)	0.54	0.596	1.7	0.2
upper lip center (x)	0.83	0.412	4.4	0.7
center of chin (x)	1.71	0.097	5.3	1.6

**Table 3.1 Magnitude of fluctuating asymmetries and directional asymmetry for the face stimuli of Experiment 1.** The above table shows the magnitude of fluctuating and directional asymmetries for a number easily located facial feature points. For each fluctuating asymmetry value a t-test against a hypothesised mean of zero (df=29) was performed. The values from the t-test are also shown.

Experiment 2 formed the basis of work published in *Human Evolution and Behaviour* by Perrett *et al.* (1999). Since the studies in Experiment 1 and 2 were performed, Rhodes *et al.* (1998; 2001a) have published similar results, using techniques similar to those used by Swaddle and Cuthill (1995), with the addition that the original face images were manually retouched, prior to being made symmetrical, to remove distinguishing marks that might cause the problem of double blemishes when the original and mirror images were combined. Rhodes *et al.* (2001a) have also shown that the preference for symmetry is not just present in Western cultures but appears to be a generally attractive trait being preferred by both Chinese and Japanese as well as Western subjects. The morphing technique used to make face images more symmetrical has been extended to show that bodies, when made more symmetrical are also more attractive (Tovee, Tasker, & Benson, 2000). Thus, the notion that manipulation to make humans more symmetrical increases attractiveness is widely accepted but the reasons why symmetry is preferred may still be debated.

Enquist *et al.* (1994; 1997) suggest that preferences for symmetry may be a by-product of the way that the human visual system works, rather than being related to the guiding of mate choice. People do seem to prefer symmetry in many things. The objects that human industry manufactures are generally symmetrical, apart from when function dictates otherwise. We often prefer the products of nature to be symmetrical e.g. flowers, fruit, gemstones while non-human animals tend to be more highly prized if symmetrical. This preference for symmetry, possibly learned, may just be generalized into the psychological tasks that subjects were asked to perform when judging facial images in experiments. However, high levels of symmetry do not only make faces appear more attractive, but also more healthy (Jones *et al.*, 2001; Rhodes



et al., 2001b) and there is evidence that the perceived health benefits related to being more symmetrical might be something other than the product of a 'halo effect'.

Jones *et al.* (2001) examined the interplay between the perception of health and attractiveness of faces differing in level of symmetry. If symmetry is preferred, either because of a generalization of symmetry preference for other objects or as a by-product of the way the human visual system worked, then it might be expected that a preference for symmetrical faces over less symmetrical versions would be mediated by a general attractiveness halo effect (see Feingold, 1992; Langlois et al., 2000, for meta-analysis and review of the halo effect). Instead, Jones *et al.* found that the effect of symmetry on ratings of attractiveness seems to be mediated by factors that are related to subjects' ratings of the individuals' health.

If symmetry is important in judgements of attractiveness for evolutionary reasons, then it might be expected that individuals who are more attractive might be more demanding of symmetry. This effect might be expected to be present only for the faces of individuals of the opposite sex to the rater. The next experiment tested this hypothesis with the facial image stimuli from Experiment 2 in which symmetry was manipulated. Self-perceived attractiveness was used as a measure of attractiveness as it has been shown to be a good measure of attractiveness (Feingold, 1988).

### ***Experiment 3: The influence of self perceived attractiveness on selectivity***

#### **Aims / hypothesis**

The aim of this experiment was to test whether female subjects who think themselves more attractive will show a greater preference, i.e. would be more

selective, than female subjects who feel themselves less attractive, for *symmetrical* over *normal* versions of opposite sex face images.

## Methods

### Subjects

94 female subjects (average age 21 years) completed the experiment.

### Presentation

After completing a questionnaire that included a question on self-rated attractiveness, subjects went on to perform an experiment in which they were presented with the pairs of faces used in Experiment 2. The face pairs were presented

**Which face is the most attractive?**



**Figure 3.6. Stimulus presentation interface as used in experiments 3 and 4**

**Please click the face which you feel is most attractive.**

in random order with male and female trials randomly interspersed on the subject's computer across the web, via a JavaScript presentation program (illustrated in Figure 3.6). The use of the web means that there may be variations in how the test appeared to different subjects. Although, in order to run the test, all subjects needed to be online with and using a recent version of Netscape Navigator or Microsoft Internet Explorer. These conditions would suggest to us that the great majority of individuals who were

able to perform the experiment would have a screen set to high colour (16-bit) and thus the experimental stimuli would appear to be very similar to different subjects.

## Results

To test whether subjects who think themselves more attractive are more selective for opposite sex symmetrical faces, but not same sex faces, correlations were performed on the relationship between subjects' self reported attractiveness, the proportion of symmetrical male face stimuli chosen as being most attractive and the proportion of symmetrical female face stimuli chosen as being most attractive. There was a significant correlation between self-reported attractiveness and the proportion of symmetrical male stimuli chosen ( $N=97$   $r=0.242$   $p=0.014$ ), but no significant correlation between attractiveness and the proportion of symmetrical female stimuli chosen ( $N=97$   $r=0.014$   $p=0.89$ ).

Subjects who were more selective for symmetry in opposite sex faces were also more selective for symmetry in same sex faces since there was a significant correlation between the proportions of symmetrical male and female stimuli chosen ( $N=97$   $r=0.253$   $p=0.012$ ). It is possible that individuals who are more selective in terms of symmetry generally may also think themselves more attractive and be driving the effect. To ascertain whether this was the case, a partial correlation was performed between self-perceived attractiveness and the proportion of symmetrical male faces chosen as more attractive, with the proportion of symmetrical female faces chosen partialled out. The resulting correlation was significant ( $r=0.246$   $df=94$   $p=0.016$ ), suggesting that the relationship between self-perceived attractiveness and selectivity for opposite sex (male) faces is not related to a general liking for symmetry.

		Female faces		Male faces	
		Average	SD	Average	SD
Self rated attractiveness	1	8.67	1.53	7.33	1.53
	2	6.95	1.61	7.58	1.68
	3	7.43	2.25	7.66	2.14
	4	7.45	2.65	9.05	1.76

**Table 3.2. Mean and standard deviation (SD) of the number of faces symmetrical faces selected by female subject who rated their attractiveness as either 1, 2, 3 or 4.**

To enable data to be understood more easily mean and standard deviations for each level of symmetry preference are shown in table 3.2.

## **Preliminary discussion**

The results of Experiment 3 suggest those female subjects who think themselves more attractive are more selective or demanding regarding male face symmetry. The effect appears to be specific to the sex of the face since there was no relationship between self-perceived attractiveness and selectivity for female face symmetry. A positive correlation between selectivity for symmetry in male faces and selectivity for symmetry in female faces was also found. This correlation might reflect some subjects generally preferring symmetry in faces, being attentive to the task in general or being more attentive to the specific symmetry cues than other subjects. The partialling out of a general preference for symmetry had no discernable effect on the relationship between self-perceived attractiveness and selectivity for symmetrical male faces. These results are inconsistent with the view that the preference for symmetry is just a by-product of the way that the visual system works (Enquist & Arak, 1994) since the preference depends on the condition of the observer, and the effect found here is only for opposite sex face judgements. Instead the results are consistent with the view that the importance of symmetry in attractiveness judgements reflects adaptations to facilitate mate selection.

Thus, it appears that female subjects who think themselves more attractive show a greater preference (than female subjects who feel themselves less attractive)

for symmetrical over less symmetrical versions of male face images. Modulation of visual preference has also been found for judgements of the level of masculinity most preferred in male faces, depending upon how a woman feels about her own attractiveness (Little, Burt, Penton-Voak, & Perrett, 2001) and upon whether she is in a phase of her menstrual cycle in which she is likely to conceive (Johnston, Hagel, Franklin, Fink, & Grammer, 2001; Penton-Voak et al., 1999). Analogous modulations of preference have been found in the domain of smell, with women, who are at the most fertile stage of their menstrual cycle, showing a greater preference for the scent of symmetrical males (Gangestad & Thornhill, 1998) and a greater preference for the smell of individuals with dissimilar immuno-histocompatibility complexes to their own<sup>21</sup> (Penn & Potts, 1999).

The data for Experiment 3 has also been reported using a different statistical analysis as part of a paper by Little, Burt, Penton-Voak & Perrett (2001). This paper examined the influence of self-perceived attractiveness of female subjects in the light of findings in sticklebacks (*Gasterosteus aculeatus*) and guppies (*Poecilia reticulata*). Female sticklebacks generally prefer to mate with male sticklebacks that signal their health with their bright red throats (Millinski & Bakker, 1990), but females who are in poor condition show a diminished preference (Bakker, Kunzler, & Mazzi, 1999). Lopez (1999) examined female guppies' mating behaviour and found a comparable result, that parasitized female guppies did not prefer the more showy males. Little *et al.* (2001) suggest that the differences in preference between high and low attractive females may be due either to the problem of low attractive females being less able to extract parental investment from a high attractive mate than higher quality females or

---

<sup>21</sup> This is interpreted as a method of conferring greater immune system diversity upon procreates.

because for high attractive females, the cost of loss of investment from a mate may be lower as they may have better resources (including health) and may find it easier to extract investment from other individuals (including other mates).

Physical attractiveness is, however, just one aspect of mate condition.

Experiment 4 explores another measure that might give some indication of how well a female individual is coping with the stresses and strains of every day life. Females who are coping better might be expected to prefer cues to a higher quality mate and therefore more symmetrical male faces. In a similar fashion as postulated above for females who rate themselves as being more attractive, females who show signs of coping well should be more able to extract relationship investment from males and be more able to deal with a lack of investment from high attractive males. As a measure of how well subjects feel they were coping with everyday life Beck's Depression Inventory (BDI, Beck, Ward, Mendelson, Mock, & Erbaugh, 1961) was used. It is noted that the use of this as a measure of depression, particularly when used with university students, has been criticised and has been suggested to indicate a general malaise as BDI score correlates highly with measures of anxiety rather than being a measure of depression (for review see Gotlib, 1984). The BDI score is interpreted indication of general malaise, low self esteem and lack of ability to cope rather than as a measure of depression.

### ***Experiment 4: Malaise and selectivity***

#### **Aims / hypothesis**

The aim of the experiment was to test whether female subjects whose self reported feelings suggest that they are coping better, as indicated by a lower BDI



score, show greater preference (than female subjects with higher BDI scores) for *symmetrical* over *normal* versions of face images.

## **Methods**

### **Subjects**

33 female subjects (average age 20.5,  $sd=0.43$ ,  $max=30$ ,  $min=17$ ) who had not taken part in previous attractiveness and symmetry experiments completed the experiment.

### **Presentation**

Female subjects completed the test as presented in Experiment 3. They were then recalled a few weeks later to complete the BDI questionnaire.

### **BDI score**

Subjects' BDI score ranged between 0 and 26 with a mean of 7.9,  $sd=1.13$ . The BDI score is a tool for the quantification of depression in clinically diagnosed depressed individuals. Lower scores indicate less depression, with BDI scores above 10 widely being interpreted as representing depressed individuals<sup>22</sup>. It should be noted, however, that high BDI scores in a student population who have not been clinically diagnosed with depression should not be interpreted as a simple indication

---

<sup>22</sup> In a later chapter a BDI cut-off of 15 is used to differentiate between depressed/disphoric non-students since recent work has suggested that using high cut-offs (above 16 or above 21) results in a group who are similar in symptom severity to clinically diagnosed depressed individuals (Cox, Enns, & Larsen, 2001).

of depression (Gotlib, 1984). Here high BDI scores are interpreted as a reflection of a more general malaise.

## Results

Female subjects with higher BDI scores were significantly more selective for male face symmetry ( $N=33$   $r=0.431$   $p<0.05$ ). There was no significant relationship between BDI score and selectivity for female face stimuli ( $N=33$   $r=-0.017$   $p=0.942$ ). To remove any general effect of selectivity for symmetry, the proportion of symmetrical female faces chosen by each subject was partialled out of the correlation between preference for symmetry in male faces and BDI. This partialling out did not diminish the significance of the correlation ( $df=30$   $r=0.445$   $p<0.05$ )<sup>23</sup>. Descriptive statistics for participants with low and high BDI are shown in Table 3.3.

BDI score	Female faces		Male faces	
	Average	SD	Average	SD
Low	7.49	1.86	7.12	1.62
High	7.64	1.70	8.26	2.01

**Table 3.3. Average and standard deviation of the number of symmetrical male and female faces by subjects with low BDI (BDI  $\leq 10$ ,  $N=25$ , average BDI= 4.9, sd= 3.3) and high BDI (BDI  $>10$ ,  $N=8$ , average BDI= 17.2, sd= 5.0).**

The Beck Depression Inventory Manual (Beck & Steer, 1987) examines the answers given to individual questions that make up the BDI questionnaire. To explore the reasons for the above effect, the relationship between the individual questions that make up the BDI and female selectivity for male symmetry was examined. A positive correlation was apparent between selectivity for male symmetry and the questions related to: sadness (BDI question 1 {see Chapter 3 Appendix};  $N=33$   $r=0.384$   $p=0.027$ ), punishment (BDI question 6;  $N=33$   $r=0.489$   $p=0.004$ ) and crying (BDI

---

<sup>23</sup> The slightly higher 'r' value may be a result of the partial correlation, unlike the original correlation, being parametric.

question 10;  $N=33$   $r=0.450$   $p=0.009$ ) but no other significant effects for any single questions. For the BDI questions see Chapter 3 Appendix.

## **Preliminary discussion**

Contrary to expectations there was a positive correlation between BDI and female selectivity for symmetry in male (but not female) faces. In other words, female subjects with a higher degree of malaise (as indicated by higher BDI score) chose a greater proportion of symmetrical face images. There was no similar correlation for selectivity for symmetrical female faces. Study of the individual questions that make up the BDI revealed a high correlation between selectivity and the answers to the question for feeling of being punished and a smaller correlation between selectivity and the questions relating to sadness and crying.

The reasons for this correlation between higher BDI and selectivity for male faces are unclear. Subjects that are more selective in terms of the symmetry of their partners might be more likely to meet with disappointment and feel punished within their actual relationships, possibly because females who are very selective for symmetry might “be setting their sights too high” and risk being jilted by males with greater symmetry (and attractiveness).

An alternative interpretation is that the BDI test is not the appropriate tool for investigating the concepts that are examined here. Its use, especially in student populations, has been criticised and it is debated whether non-diagnosed individuals, particularly students, who are run on the BDI are comparable to clinically depressed subjects (Gotlib, 1984). In this experiment, the BDI test is not being used as a means of quantifying depression, but as an index of how well individuals are coping. It is felt that such an interpretation of BDI scores is not at odds with the criticisms levelled its the use in non-depression diagnosed student populations.

## ***General chapter discussion***

This chapter has examined a long running set of experiments that started by showing, and then confirming using more conservative methods, that facial symmetry was preferred in humans. Before the first two experiments were performed, the literature suggested that symmetry was widely preferred in non-human animals for mate choice, but not in humans, where a preference for asymmetrical faces as being more attractive existed. Thus, the first two experiments presented here, together with experiments performed around the same time by Rhodes *et al.* (1998), have brought the human literature in line with the animal literature by showing that, like other animals, humans find symmetry attractive in potential mates.

Symmetry is preferred in many different areas of our lives, a preference for symmetrical faces may reflect a general preference, or may be the result of how our visual system functions (pattern recognition systems may find symmetrical objects more easy to recognise leading to their being perceived as more attractive, Enquist & Johnstone, 1997). The results of Experiment 3, and other studies, suggest that symmetry preferences are more than simply a by-product of the workings of the visual system or a general preference. Jones *et al.* (2001) show attractiveness judgements of facial stimuli that vary in degree of symmetry are better explained by perceptions related to apparent health rather than to an attractiveness halo effect. The work presented here shows that the attractiveness of the viewer is influential in how selective female individuals are when it comes to symmetry of male faces. This effect was specific to opposite gender faces, since the effect was not related to how selective females are for symmetry when presented with female faces. This opposite sex specificity is significant, since if the mechanisms that make symmetry salient are

general, rather than being related to mate choice, then it would be expected that similar effects should be equally present for choices of female as for male faces.

Experiment 4 proceeded to explore the relationship between BDI score and selectivity for symmetry in female observers. It was expected that female observers who scored lower on the BDI scale would be more selective for male face symmetry. Contrary to expectations it was found that female observers that were higher in BDI score were more selective. This result might be explained in terms of being selective relating to other personality traits, such as being overly critical, that might result in a high BDI score. More work with other indexes of coping and psychological well-being is required to understand this relationship.

The work that is presented in this chapter started some time ago but the investigation of the symmetry in mate choice is still a 'hot topic', large amounts of research work being published (currently the Royal Society of London alone is contributing 3 articles a month on symmetry and its relationship to mate selection in non-human animals). Research is being performed to examine factors in individuals that influence their preference for symmetry (including Experiment 3 and 4 presented here, see also Jones et al., 2001; Little et al., 2001). Other studies are probing the amount of influence that symmetry has over judgements of attractiveness, compared to other cues (Kowner, 2001; Penton-Voak et al., 2001; Scheib et al., 1999). The exploration of symmetry is a topic that is relevant to the behavioural sciences both now and probably for some time to come.

When the work contained in this chapter was started, the methods of graphic manipulation used (both photographs and computer graphics) indicated that symmetry was not preferred in humans. Here by adaptation of computer graphics techniques appropriately to manipulate facial symmetry a paradigm has been devised that has

utility in the study of facial attractiveness and has shown that in humans, as in other animals, symmetry is preferred. Furthermore these graphic techniques have been shown to be of value in the study of individual differences in the perception of symmetry as attractive.



## **Chapter 4: Manipulation of visual cues to perceived health in faces.**

## ***Introduction***

Attractiveness is highly important in social interaction (Dion, Berscheid, & Walster, 1972; Ritter & Langlois, 1988) and outward indications of health and reproductive potential are a major component of attractiveness (Gangestad, Thornhill, & Yeo, 1994). Currently there is great interest in the field of facial attractiveness, mainly in connection to evolutionary fitness (for review see Simpson, Gangestad, Christensen, & Leck, 1999; Thornhill & Gangestad, 1999). Many of these studies of facial attractiveness have concentrated on the structure of the face and elements of the face that may be relatively stable over the adult life. For example, the study of facial asymmetry (Mealey, Bridgestock, & Townsend, 1999; Perrett et al., 1999; Rhodes, Proffitt, Grady, & Sumich, 1998; Swaddle & Cuthill, 1995) assume poor health, particularly during growth, leads to increased fluctuating asymmetry (Møller, 1990). Thus, fluctuating asymmetry may be used in mate selection by many species as a visual indicator of disease resistance and good genes (Gangestad et al., 1994; Manning & Chamberlain, 1993; Thornhill & Gangestad, 1999).

The above studies of attractiveness investigate factors that are fairly constant after growth is completed. The current study, on the other hand, investigates the appearance of skin, an organ that changes dynamically from day to day (e.g. because of changes in hydration level, current state of health) as well as over the lifetime of an individual. Skin condition may also relate to apparent health and indirectly to attractiveness. If signs of resistance to disease underlie facial attractiveness then indicators of health, other than the structural aspects of the face should be important.

The faces of individuals differ in many ways, as shown by the many Principal Components that are needed for an adequate representation of the identity of a human face image from within a set (for example see Hancock, Burton, & Bruce, 1996;

O'Toole, Deffenbacher, Valentine, & Abdi, 1994). In a similar way both facial health and attractiveness are likely to be manifest in many different ways. This means that one of the major problems in studying perceptual health is specifying the different factors. The different health factors used here are based on a study in which health descriptions of faces were elicited from groups of observers (Hawkins, Perrett, Burt, Rowland, & Murahata, 1999). An analysis was performed to specify compactly why different faces were thought to be more or less healthy. This analysis returned 3 factors that corresponded closely to the 3 descriptors that were chosen for use in this study.

The first health descriptor chosen relates to ratings of skin wrinkles. As individuals age, their skin becomes less elastic and more wrinkled. Generally health (and beauty) is associated with youth (Henss, 1991), so faces with skin that looks more lined and wrinkled may be seen as less healthy.

A second descriptor of health that is manifest within the face is often related to the bodies' response to bacterial infection of the skin surface. This is associated with a variety of visual changes, including acne or spots, pitting or scarring with recurrent acne, a widening of skin pores and a general inflammation which is manifest as a reddening of the affected areas. Such inflammation occurs in topologically specific regions of the face (generally the forehead, nose, chin).

A third descriptor of apparent health concerns the classic 'rosy cheek' appearance reflecting superficial blood flow particularly to the cheek region. The opposite of such a healthy complexion is manifest when the blood drains away from the cheeks triggering descriptions such as pallid, ashen, yellow or sickly.

These three descriptors of facial health were studied and are referred to below as Lines-Wrinkles, Inflammation and Glow-Colour respectively. The procedures that

were used to investigate these descriptors are derived from the work of Galton (1878) who photographically superimposed face images matched for eye position onto photosensitive paper. His method gives rise to face images which, although blurred, incorporate many of the similarities of the group of faces. The computer graphic method of 'warping' permits the position of many features within faces to be matched, so that a clearer blend image can be formed (Benson & Perrett, 1993). Despite these advances the position of some details (e.g. fine wrinkles) cannot be matched before the images are blended resulting in a loss of fine grain features. Even so, the technique of making blends used here captures many cues and blends have been shown to capture at least 75% of the information related to age (Burt & Perrett, 1995).

Previous studies have successfully extracted facial traits related to attractiveness (Perrett, May, & Yoshikawa, 1994), age (Burt & Perrett, 1995) and gender (Rowland & Perrett, 1995) by blending together groups of faces. Some traits, however, may simply not be in a visible form in face images. For example, Galton (1878) wanted to compose a portrait of the facial traits related to criminality by superimposing many faces of criminals. However, he found that the images he had created of the average criminal were "much better looking than those of the components" (Galton, 1878 p.97). He was unable to extract facial cues to criminality from faces probably because they do not exist (though see Berry & McArthur, 1986). The same may be true of health: there may not be any consistent facial cues that relate to health.

If information can be captured in facial blend images, then it may be possible to use this information to alter the appearance of individual facial images. Facial blends allow the face image of an individual to be moved, or transformed from having

the characteristics of one group of faces to having the characteristics of another group of faces. For example the colour and shape difference between a blend of old faces and a blend of young faces can be used to change the apparent age of an individual (Burt & Perrett, 1995). Alternatively the gender of an individual may be changed by applying the difference between an average male face and an average female face (Rowland & Perrett, 1995; Yamada, Chiba, Tsuda, Maiya, & Harashima, 1992). Here the impact of attempts to transform the perceived health of faces using the same principles is assessed.

In this chapter, 2 experiments are presented. The aim of the first experiment was to check the validity of the descriptors captured, i.e. that a blend of individuals rated high on a descriptor receives higher ratings on that descriptor than a blend of low on the descriptor.

The second experiment was to establish whether colour transforms would be used by subjects to maximise the perceived health of face images. Many graphical manipulations of face images are possible but most manipulations will result in a less healthy appearance. To show that a health transform is manipulating a health descriptor it is therefore important to demonstrate that it can be used to make face images appear healthier. Five naïve subjects will therefore be tested on whether, when asked to make faces look optimally healthy they choose the original, unchanged face image, or choose to alter the face by using a colour transform related to one of the health descriptors. Additionally, one might expect that images of faces originally rated as low on a health descriptor might have greater degrees of health transform applied to them in order to make them look healthier than images originally rated high.

## **Experiment 1**

### **Methods**

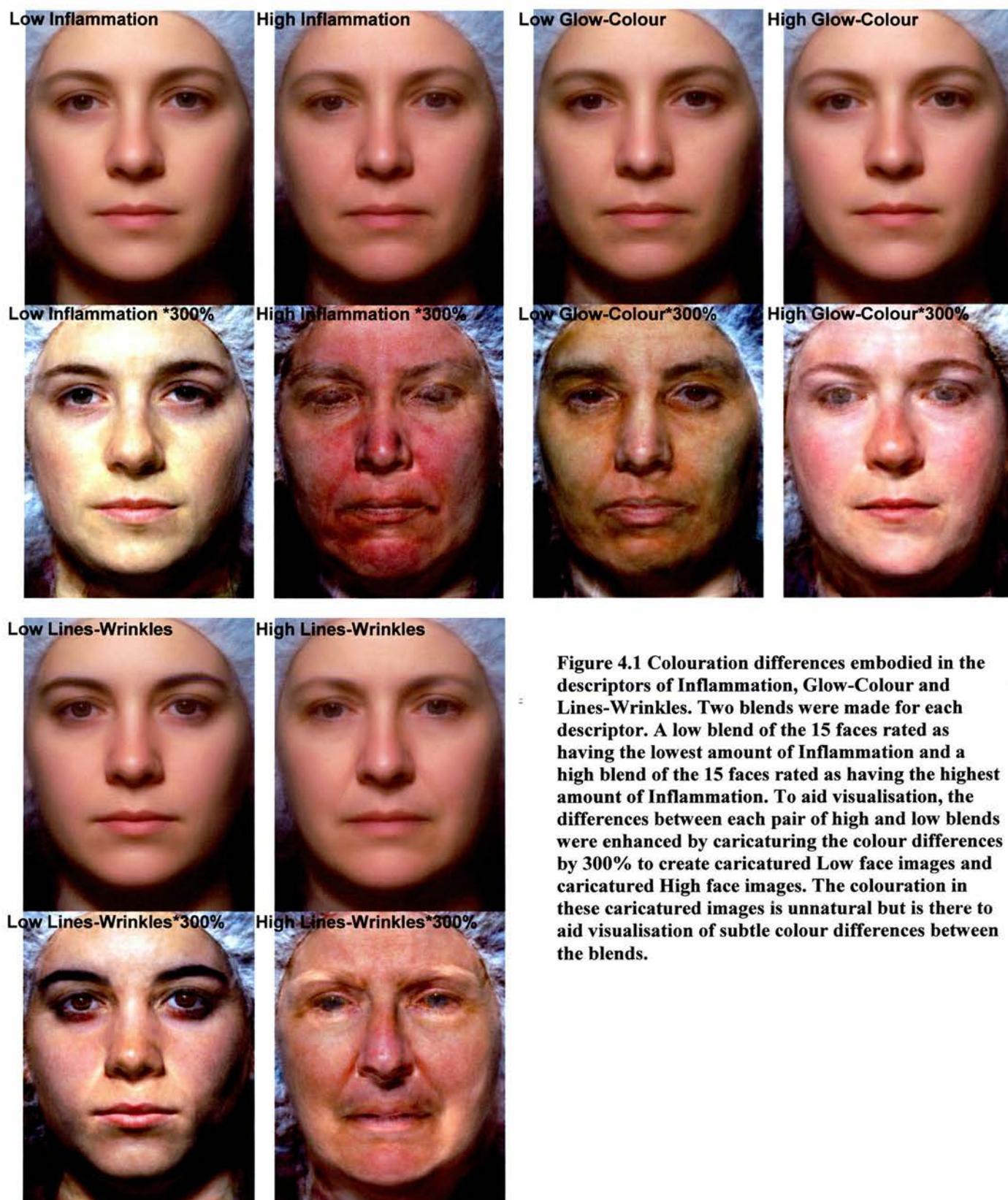
#### **Collection of health ratings**

Six skin condition raters based with Unilever Skin Research at Edgewater, USA rated the health of 62 female faces (aged between 18 and 59 with a mean age of 36 years) on multiple health descriptors. The skin condition raters evaluated each face under standard lighting conditions while the person stood in front of them. For each of the 3 perceived facial health descriptors the faces were rated on a 5-point scale. After being rated, the faces were then photographed under standard lighting conditions, using daylight-balanced flashlight with a hairnet to cover their hair and a black shawl covering their neck downwards. Photographs were transferred from 35mm-slide film onto Kodak Photo CD. Further processing was performed at Kodak Photo CD size 4 (1024 by 1536 pixels), at 24-bit colour and an average eye separation of 279 pixels.

#### **Delineation and computation of blends**

The shape of each face was registered by manually positioning 178 landmark points onto matching features on each face image (e.g. one over the left lip corner, another at the centre of the top of the lip, etc.). The average face shape for the population of 62 faces was then calculated as the average x, y co-ordinate value for each feature point. The faces were then warped into the average shape to allow two blends (Benson & Perrett, 1991; Rowland & Perrett, 1995) to be made for each health descriptor: a Low blend made by averaging together the colour of corresponding pixels for the 15 faces rated lowest on the descriptor, and a High blend made by averaging together the 15 faces rated highest on the descriptor. The areas outside the





**Figure 4.1** Colouration differences embodied in the descriptors of Inflammation, Glow-Colour and Lines-Wrinkles. Two blends were made for each descriptor. A low blend of the 15 faces rated as having the lowest amount of Inflammation and a high blend of the 15 faces rated as having the highest amount of Inflammation. To aid visualisation, the differences between each pair of high and low blends were enhanced by caricaturing the colour differences by 300% to create caricatured Low face images and caricatured High face images. The colouration in these caricatured images is unnatural but is there to aid visualisation of subtle colour differences between the blends.

face were then masked in black (for examples of blends see Figure 4.1). Thus 6 blends were made: a High and a Low blend for Inflammation, a High and a Low blend for

Glow-Colour and a High and a Low blend for Lines-Wrinkles.

## **Subjects**

18 subjects (10 female and 8 male, mean age = 26 years) took part in the experiment.

## **Procedure**

The 6 blended images were presented in a random order via computer to subjects. Subjects rated each stimulus on the relevant descriptor from which they were derived and on health on a 7-point scale running from “Not at all” to “Very highly”.

## **Results**

Subjects rated the High blend as being higher on the appropriate descriptors than the Low blend for Inflammation (Wilcoxon Signed Ranks Test comparing subjects' ratings of the High and Low blends  $n=18$ ,  $Z=-2.37$ ,  $p<0.05$ ), Glow-Colour ( $n=18$ ,  $Z=-3.54$ ,  $p<0.01$ ) and Lines-Wrinkles ( $n=18$ ,  $Z=-3.65$ ,  $p<0.05$ ). The health ratings of the High blend was significantly higher than that for the Low blend for Lines-Wrinkles ( $n=18$ ,  $Z=-2.23$ ,  $p<0.05$ ) but there was not difference in health ratings between the High and Low blends for Inflammation ( $n=18$ ,  $Z=-0.88$ ,  $p=0.38$ ) or Glow-Colour ( $n=18$ ,  $Z=-1.51$ ,  $p=0.13$ ).

## **Discussion**

Subjects rated the Low blend (made from face images of people originally rated low on the descriptor) as possessing less of the relevant descriptor (Inflammation, Glow-Colour or Lines-Wrinkles) than the High blend (made from face images of people originally rated high on the descriptor). This shows the blend images differ on the relevant descriptors. For ratings of health of the different blends there

was only a difference between the High and Low blends for Lines-Wrinkles. The lack of difference between Low and High blends for Inflammation and Glow-Colour may be because this type of test using a ratings scale is not sensitive to pick up subtle differences in the perceived health of face images.

## ***Experiment 2***

### **Method**

#### **Selection of faces and generation of trials**

For each of the descriptors (Inflammation, Glow-Colour and Lines-Wrinkles) 9 faces were chosen: 3 originally rated highly on the descriptor, 3 originally with mid level ratings and 3 originally rated low. This gave 27 face images in all. Each face image was used to construct a trial in which subjects were able to change the image by adding and subtracting a pre-chosen colour transform. Subjects added and subtracted the transform to the face image during the trial by moving the mouse from left to right. Mouse movement resulted in a real time face image change achieved by using a computer technique called 'alpha-blending' to alter the transparency between a pair of 'end-point' images such that one image gradually fades into the other as the computer mouse is moved from the left to the right of the screen. One of the end-point images in the pair was the original face image transformed to have 200% (+an offset) more of the difference between the High and Low blend of one descriptor and the other end-point image was transformed to have 200% (-an offset) less of the chosen descriptor. So in a trial moving the mouse to the far left of the screen would result in one end-point image being displayed where as moving the mouse to the other side of the screen would result in the other endpoint image being displayed. Intermediate

points on the screen would result in the display of intermediate amounts of transform. A random offset of  $\pm 50\%$  of the applied colour transform was used to ensure that the original image did not systematically occur when the computer mouse cursor was in the centre of the screen.

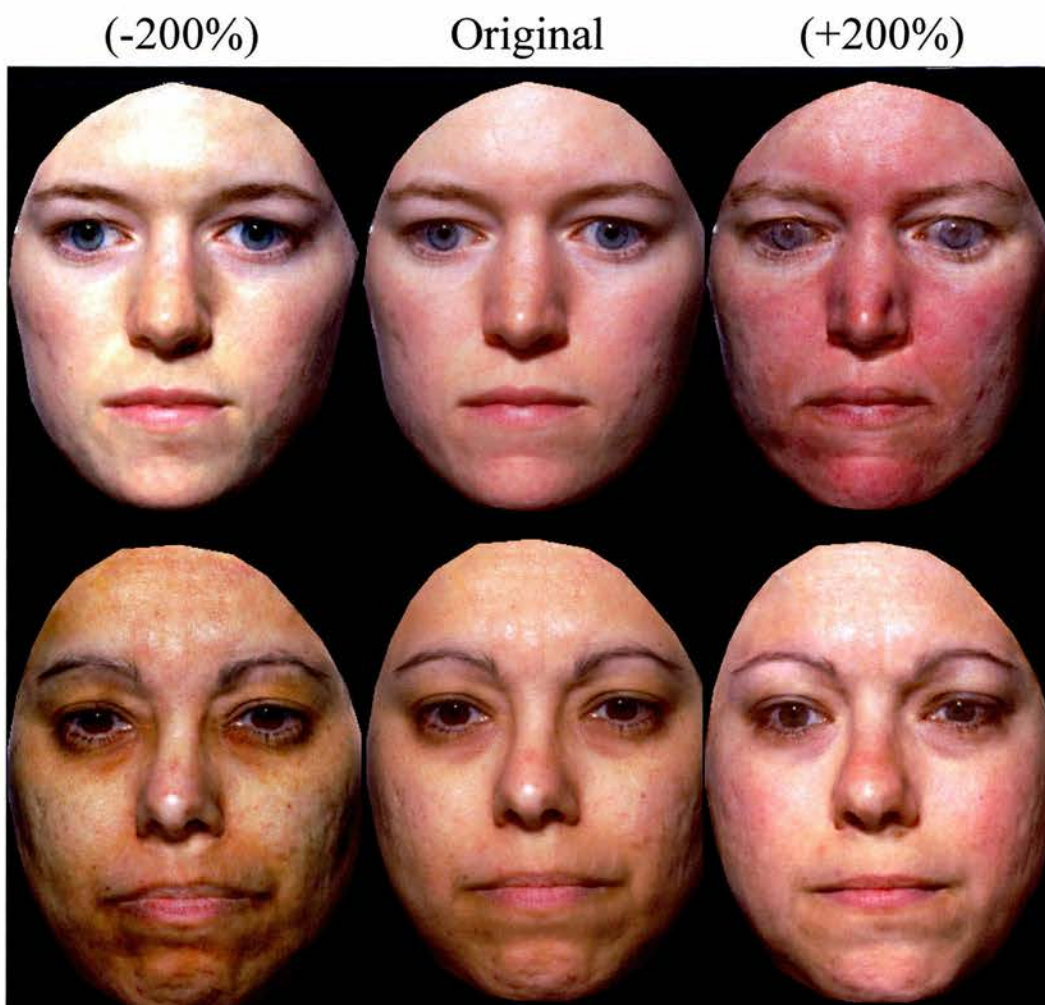
End-point images were produced using colour transforms. In this technique (Burt & Perrett, 1995; Rowland & Perrett, 1995), two blend images are warped into the same shape as the original image. For each pixel of the original image the colour difference between each of the 2 blends for the corresponding pixels is calculated, multiplied by the amount of transform required and added to the original pixel value. To make the end-point images this value is  $200\% + \text{offset}$  for one of the image and  $-200\% + \text{offset}$  for the other image.



A pair of practice trials was manufactured using 2 additional individuals who were picked randomly from the set. A health transform related to a descriptor called ‘skin dryness’ that produced changes that were judged to be visually unlike the changes produced by the other three transforms used here was applied to the images. For all trials the hair and other regions outside the face were masked in black. Examples of the end-points of the stimuli presented to subjects can be seen in Figure 4.2.

## Procedure

After two practice trials to permit the subjects to become familiar with the procedure, subjects started the test trials. Subjects were instructed to move the mouse



**Figure 4.2** Example end point and original face images for the Inflammation (top row) and the Glow Colour (bottom row) health descriptors. Images on the left have been transformed to subtract twice the colouration difference between the High and Low blends from the original image. Images on the right have been transformed to add had twice the colouration difference between the High and Low blends to the original image.

left and right to make the displayed face look as healthy as possible. These horizontal mouse movements produced a real time colour transform of the faces. When the subjects had made the face look maximally healthy they were asked to press the space bar. Pressing the space bar ended the trial and started the next trial. Subjects proceeded through the 27 randomly ordered trials, seeing each face stimulus once. The trials were presented to the subjects on a 20-inch Silicon Graphics monitor in 16bit colour, at full screen size.

## **Subjects**

11 subjects, 6 female and 5 male (aged between 22 and 35 years) performed the experiment.

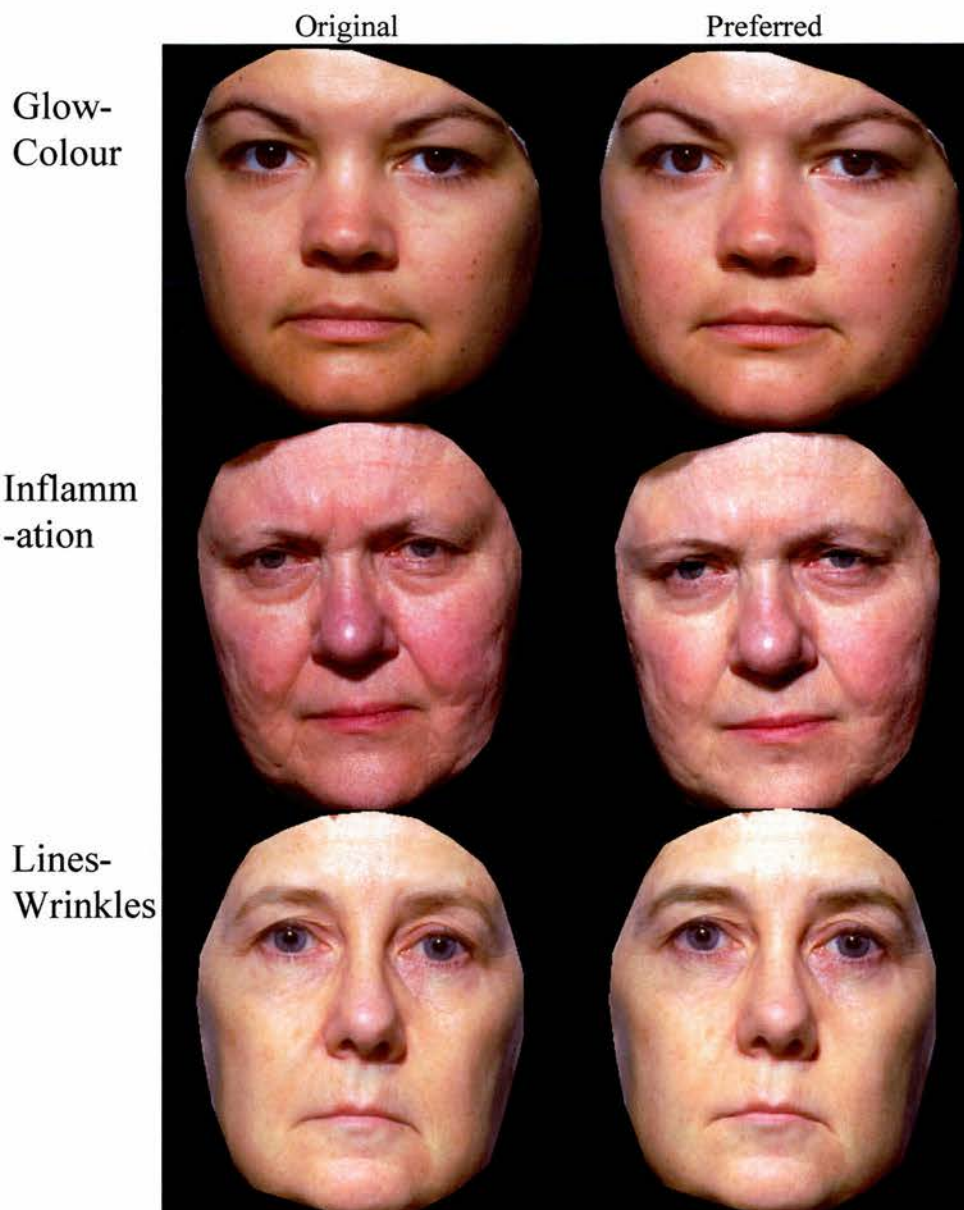


## Results

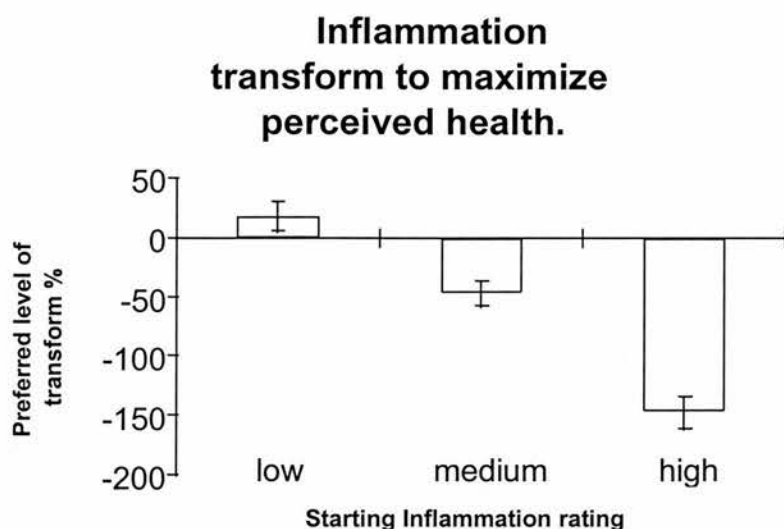
Examples of the average colour differences preferred by subjects can be seen in Figure 4.3. Figures 4.4 to 4.6 summarise the data from Inflammation, Glow-Colour and Lines-Wrinkles transformations respectively.

### Inflammation

To make faces look maximally healthy subjects chose to transform face images by a negative amount of Inflammation transform (t-test against a hypothesised



**Figure 4.3 Example subject preferences. Subjects on average chose to transform the above faces by adding 100% of the Glow-Colour transform to face 1, subtracting 179% of the Inflammation transform to face 2, and subtracting 79% of the Lines-Wrinkles transform from face 3 when selecting the most healthy looking appearances.**



**Figure 4.4 Preferred amount of Inflammation transform ( $\pm 1$  SE) for faces with different starting levels of Inflammation. One hundred percent of transform is equivalent to transforming the original images by the colour difference between the High and Low Inflammation blends.**

mean of zero transform,  $t_{10}=-4.1, p<0.001$ ). Further analysis by one way ANOVA (starting amount of Inflammation, three levels; subjects, 11 levels random factor) revealed a significant main effect of starting level of health rating for Inflammation ( $F_{2,10}=53.7, p<0.001$ ). This main effect reflected the subjects applying different levels of Inflammation transform to maximize the facial image's perceived health depending on the original rating of the face for Inflammation. The amount of Inflammation transform applied by subjects was greater for faces with a high starting rating for Inflammation than for faces that originally had a medium starting rating (PLSD post hoc test, high original rating group < medium;  $p<0.05$ ). Similarly, the transform applied to faces starting with a medium Inflammation rating was greater than that for faces starting with a low rating for Inflammation (PLSD; medium original rating group < low;  $p<0.05$ ). For the faces with a low starting rating of Inflammation, the amount of transform applied was not different from zero (t-test against a null hypothesis mean of zero transform;  $t_{10}=1.47, p=0.17$ ).

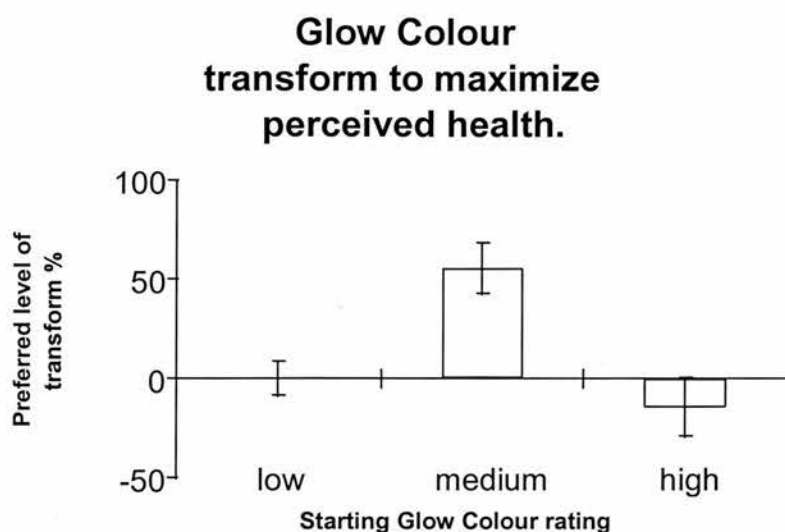


Figure 4.5 Preferred amount of Glow-Colour transform ( $\pm 1$  SE) for faces with different starting levels of Glow-Colour. One hundred percent of transform is equivalent to transforming the original images by the colour difference between the High and Low Glow-Colour blends.

## Glow colour

Overall subjects chose to transform faces by a positive amount of Glow-Colour transform which was significantly greater than expected by the null hypothesis with an expected mean of zero transform ( $t$ -test,  $t_{10}=2.5$ ,  $p=0.03$ ). The effect of starting rating of Glow-Colour on the preferred amount of Glow-Colour transform applied by subjects was assessed by one-way ANOVA (starting amount of Glow-Colour, 3 levels; subjects, 11 levels random factor). There was a significant main effect of starting level of Glow-Colour ( $F_{2,10}=12.1$ ,  $p<0.001$ ). Subjects had a significantly greater tendency to increase the level of Glow-Colour added by transform to faces which had originally started out with a medium ratings of Glow-Colour compared to low ratings Inflammation (PLSD; medium original rating group  $>$  low;  $p<0.05$ ) or a high (PLSD; high original rating group  $<$  medium;  $p<0.05$ ) starting levels of Glow-Colour. There was no significant level of transform applied to change the faces that started with low Glow-Colour ratings ( $t_{10}=0.00$ ,  $p=1.00$ ).

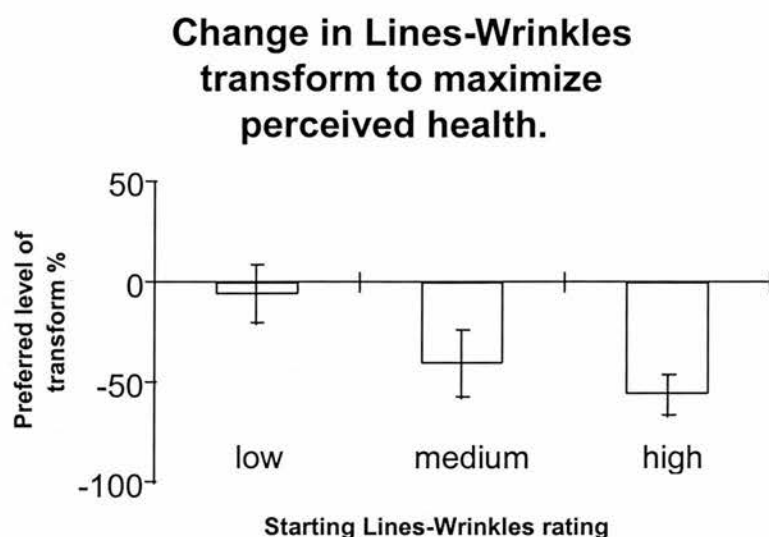


Figure 4.6 Preferred amount of Lines-Wrinkles transform ( $\pm 1$  SE) for faces with different starting levels of Lines-Wrinkles. One hundred percent of transform is equivalent to transforming the original images by the colour difference between the High and Low Lines-Wrinkles blends.

## Lines-Wrinkles

Overall subjects chose to transform the face images with a negative amount of Lines-Wrinkles transform which was significantly different from zero transform expected by the null hypothesis ( $t_{10}=-7.5, p<0.005$ ). The effect of starting rating of Lines-Wrinkles on the preferred amount of Lines-Wrinkles transform applied by subjects was assessed by one-way ANOVA (starting amount of Lines-Wrinkles, 3 levels; subjects, 11 levels random factor). There was a significant main effect of starting level of Lines-Wrinkles ( $F_{2,10}=7.1, p<0.005$ ). PLSD post hoc Tests, revealed a significant difference between the transform level applied to faces with high and low original ratings for Lines-Wrinkles (PLSD; high original rating group < low;  $p<0.1$ ).

Subjects applied a greater degree of the Lines-Wrinkles transform (in a negative direction i.e. to lessen wrinkles) as the starting rating for Lines-Wrinkles on the original face increased. The amount of Lines-Wrinkles transform applied to the group of faces that started with low starting ratings of the amount of Lines-Wrinkles was not significant ( $t_{10}=-0.51, p=0.62$ ).

## ***Discussion***

The techniques of averaging and transforming presented here were successful in capturing colour changes related to the three descriptors of perceived health. When asked to make face images look maximally healthy, subjects applied an amount of transform rather than selecting randomly or selecting the original image colouration. Moreover, subjects applied the health transforms in the manner expected: subtracting traits generally thought to appear unhealthy, both skin inflammation and lines and wrinkles, but adding Glow-Colour a descriptor related to a more glowing complexion that is thought to be healthy. The blends that these health transforms were based maintained the visual cues to the three health descriptors with the High blend (made up of the 15 faces rated as highest on the descriptor) being rated as possessing more of the descriptor than the Low blend. When asked for a more general rating of health of these blends only the High and Low Lines-Wrinkles blend differed in rating. This may surprise the reader since in the second experiment subjects applied difference between the High and Low blends to enhance the perceived health of the face images. It is thought that this apparent inconsistency is due to the differences between the Low and High blends being subtle, with each health descriptor playing only a small part in the overall health judgements. So a 100% difference in ratings of Glow-Colour relates to a much smaller difference in overall perceived health so is not picked up in Experiment 1 but may be picked up using the more sensitive procedure in Experiment 2.

The amount of transform that subjects chose to apply to face images was dependent on the original rating of that face before transformation. In general, the lower the rated health of the original face, the greater the level of transform subjects applied to induce an optimal level of perceived health. So an image of a face

originally rated as being highly inflamed would be transformed to remove inflammation by a large amount. A face originally rated as having a medium level of Inflammation would have a smaller amount of inflammation transform removed from it and faces rated as having no Inflammation would not be changed, suggesting that the transformations modified the faces in natural ways.

It was expected that subjects would choose to increase the overall amount of the positive health descriptor, Glow-Colour, in face images. This was the case both overall and for faces originally rated as having a moderate amount of Glow-Colour. It had been expected that subjects would choose to add Glow-Colour transform to faces initially rated as being low in Glow-Colour but this was not the case. One explanation for this lack of preference by subjects to increase Glow-Colour in the low Glow-Colour faces may be that these faces had other skin problems unrelated to the Glow-Colour descriptor. The Glow-Colour transform would not be expected to change other aspects of apparent health. This explanation is supported by the negative correlation between Glow-Colour and some of the other skin conditions e.g. Lines-Wrinkles ( $n=62$ ,  $p<0.10$ ,  $r=-0.21$ ), and Skin Inflammation ( $n=62$ ,  $p<0.01$ ,  $r=-0.35$ ).

Transforms are not necessarily linear with respect to how they are perceived. This can be illustrated by looking at Figure 4.1. Experiment 1 demonstrated that people rate the Low Inflammation face (*a*) as more healthy than the High Inflammation face (*b*). But as is apparent by looking at Figure 4.1, when the differences between the two blends are exaggerated to extremes the 300% low Inflammation caricature (*c*) produced does not look more healthy than either of the original blend images. Thus there are non-linearities in relationship between colour and perceived of health. There are also non-linearities in the calculation of colour. Particularly when performing large amounts of transform some red, green and blue



values of pixels in the image will top or bottom out. In other words they will become greater than the maximum value allowed (255) or lower than the lowest value allowed (0) and will therefore yield a colour that is inappropriate.

As a last minor point, this study also points to commonalities between cultures in the perception of health. Studies have suggested that cultures may differ in what they find attractive (e.g. Gangestad & Buss, 1993), although, most research has pointed to commonalities in the perception of attractiveness. The results of the study are dependent on commonalities in the perception of health by the original trained health raters in the USA and the naïve subjects in the UK. This study therefore suggests that there are commonalities in the perception of health between the US and UK. This may be a reflection of the linguistic and cultural links between the two cultures although it is noted that commonality in health perception may not be constant for all cultures.

## **Chapter 5: Facial reactions to tastes.**

## ***Introduction***

This study investigates the perception and communication of hedonic value of food tastes by adults. The modern field of taste perception started in the 1970s with Jacob Steiner's work with infants (e.g. Steiner, 1974). Steiner's work suggested that infants, only a few hours old, make stereotypic facial expressions when exposed to the 4 basic tastes. Since the infants had not yet been fed, their reactions had to be innate. This view was supported later with work that found sweet and bitter tastes induced the same reactions from anencephalic and hydrocephalic infants (Steiner, 1973).

Later work (see Berridge, 2000, for review) has suggested that there may be just two types of facial expressions elicited by taste: positive and negative hedonic patterns. Sweet tastes elicit positive reactions, very bitter tasters elicit very negative patterns and other tastes including salt and sour elicit less extreme negative or positive reactions. Rosenstein and Oster (1988) found that although newborn infants make negative responses to sour, bitter, and salt, the reactions to sour (because of the accompanied lip pursing) and bitter (because of mouth gaping) could be differentiated both from each other and from reactions to salt.

One difficulty of studying infants is that it is not possible to ask them how much they enjoy the stimuli, so studies have used different methods of inferring infant's enjoyment. These have included adult ratings of the enjoyment from the stimuli infants tasted, as well as behavioural and physiological measures from the infant (Rosenstein & Oster, 1988). Hedonically positive infant reactions include, increased sucking, heart rate, and tongue movements as well as infant relaxation. These compare to increased facial grimacing, restlessness and behaviours directed at stopping ingestion when presented with hedonically negative stimuli. Such reactions are more difficult to measure than the recording of subjects' verbalisations related to

their feelings aroused by a taste and so may be more problematic in estimating a gradient of hedonic responses. The measures may also tap into different aspects of taste reactions from verbally indicated enjoyment.

Work has been performed with both children and adults to investigate reactions to odours of differing hedonic value. Olfaction is an extremely important part of the experience commonly referred to as taste, although the experience of ingesting something is different from smelling it. Soussignan and Schaal (1996) investigated social effects on children's facial reactions to different odours. Children smelt different odours with no adult present, a familiar adult present or unfamiliar adult present. They found that children as young as 5 years old were able to mask their negative reaction to smells that they did not like and did so in the presence of an unfamiliar adult. Studies of reactions to odours have also been performed with adults. Gilbert *et al.* (1987) covertly and overtly filmed lone female undergraduates smelling and pretending to smell a variety of odours (similar to cloves, roses, urine, and rancid sweat). They found that observers of these films were most accurate at detecting hedonic state of the odour-experiencing individual when that person was told to pretend and when the individual was overtly filmed smelling the odour. Observers could not guess whether the individual smelt something that was enjoyable or not by assessing the individuals filmed covertly. These results lead the authors to conclude that observers of covertly filmed subjects cannot detect information about hedonic state. Alternatively, it is possible that the lower amplitude covertly observed taster reactions may be misinterpreted by subjects because they became accustomed to the higher amplitude acted expressions in the experiment.

Facial EMG (electro-myogram) has also been used to analyse the facial reactions of adults to the presence of different odours. Jancke and Kaufmann (1994)

measured EMG responses across 6 different facial sites to odours smelled by subjects when they were by themselves or with the experimenter. Jancke and Kaufmann found that when subjects were alone there were no significant correlations between facial EMG and the subjects' rating of how pleasurable, intense or pungent the stimuli were, but when the experimenter was present, subjects showed stronger reactions over the periocular and cheek regions during experiences of pleasant odours. Jancke and Kaufmann interpret their findings in terms of evidence for the social communicative function of facial social displays rather than Steiner's early reflex based explanation that facial displays are derived from innate reactions to tastes.

The present study sets out to examine what adults communicate when they sample sweet, bitter, sour and salty tastes presented as parts of naturalistic flavours. As previous studies have found that there is greater reaction to taste in social situations, during this experiment reactions to tastes were filmed in the presence of the experimenter. The experiment aimed to assess whether observers could differentiate taste, taste concentration and hedonic responses to flavours.

## **Methods**

### **Video-capture**

Four tasters (2 female and 2 male; average age 24 years) were videoed consuming a variety of drinks that varied in concentration (low, medium and high concentrations), taste (sweet, sour, bitter and salty) and hedonic value (as perceived by the taster). The stimuli were presented concealed in disposable cups with plastic lids through which the tasters sipped the stimuli using a straw. During the video-capture the taster sat facing the experimenter who presented them with the stimuli. A

video camera (Panasonic DV110) overtly recorded the taster's reactions over the shoulder of the experimenter.

## **Taste stimuli**

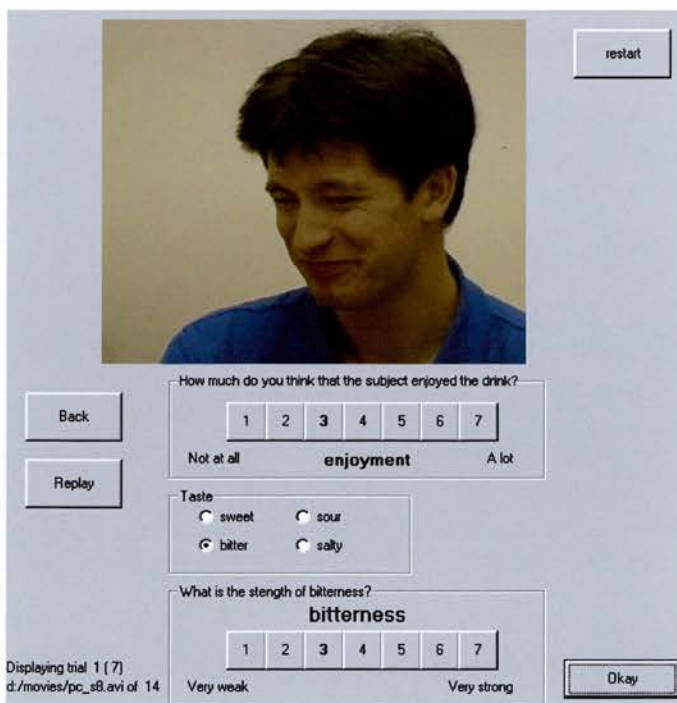
Stimulus dilutions were based on subjective quantification during a pilot experiment with 'low' being 'just detectable', and 'high' so strong that it was judged 'only just drinkable'. Medium was set as the halfway point between the 'low' and 'high' concentration.

The sour stimulus used was pure grapefruit juice (concentrations: pure; 2 parts grapefruit juice / 1 part water; 1 part grapefruit juice / 2 parts water). For the sweet stimulus, sucrose solution was used (12.5g sucrose / 150 ml water; 7.5g sucrose / 150 ml water; 2.5g sucrose / 150 ml water). Instant espresso coffee was the bitter stimulus (3.8g coffee / 200 ml water; 2.5g coffee / 200 ml water; 1.3g coffee / 200 ml water). The salt stimulus was table salt (4.8g salt / 200 ml water; 3.2g salt / 200 ml water; 1.6g salt / 200 ml water). All solutions were remade before the start of the testing for each taster using spring water and were kept before the experiment started at 4°C.

## **Procedure**

During presentation the experimenter was blind to the identity and concentration of the taste stimuli. Tasters were told that they were participating in an experiment relating to facial expressions made in response to taste stimuli. They were instructed to take small sips from the straw and to respond naturally, without concealing or exaggerating reactions. After tasting the liquid they were then asked to rate their enjoyment of the stimuli on a scale of 1-7 (1 being low enjoyment through to 7 being high). Tasters were asked to state the taste of the stimuli – sour, sweet, bitter, salt or no taste. Finally, tasters were asked to rate the concentration of stimuli on a scale of 1-7, 1 being low and 7 being high. All responses were verbalised to





**Figure 5.1.** The computer interface used by observers to rate the video stimuli of tasters' reactions while consuming liquids.

maintain visual contact between the experimenter and the taster rather than the taster looking down to make written responses. Tasters were given ad lib access to spring water and biscuits in order to maintain their palates between trials.

## Video-editing and presentation

The video recordings were edited to show just the facial responses to each stimulus, i.e. from the point when the taster stopped sucking on the straw to the moment just before they started to speak. This resulted in video clips of between 2 and 4 seconds long. A specially designed computer interface (Figure 5.1) presented the silent video clips and recorded the observer's responses to them. The video clips were presented in 4 randomly ordered blocks each containing video of one of the 4 tasters. Within each block 12 trials were presented (3 concentrations by 4 tastes).

One video clip was presented per trial. The trial started with the video clip being played. The interface then presented the question: "How much do you think that the subject enjoyed the drink?" and an enjoyment scale from 1 ("Not at all") to 7 ("A

lot”) for the observer to click their answer on. The observer was also asked to select the taste that they thought that the drink had. After they had selected the taste they then were asked by the interface to indicate on a scale how strong they thought the taste was on a scale from 1 (“very weak”) to 7 (“very strong”). After answers had been completed for all three questions and box labelled “Okay” appeared. Clicking on this box took the observer onto the next trial.

## Video-rating

Twenty-six young adult observers (14 female and 12 male) viewed computer presented video clips of the tasters consuming the stimuli and assessed the taster’s enjoyment of the drink (1-7), the taste (sweet, sour, bitter or salt), and the taste strength (on a scale 1-7).

## Results

### Hedonic value

The average observers’ score of hedonic value for each of the 48 video clips of the drinks was correlated with the score for hedonic value given by the tasters

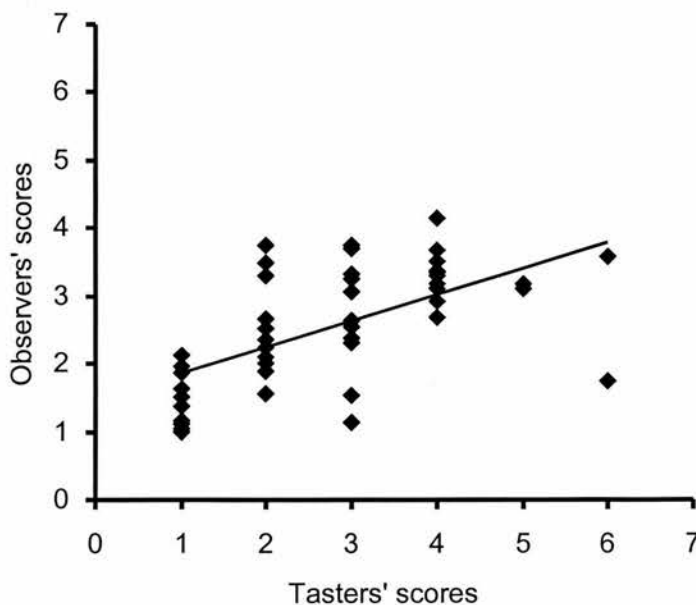


Figure 5.2 Perception of the hedonic value of taste stimuli. Correlation between tasters’ reports of enjoyment and observers’ judgments of tasters’ enjoyment.

(Spearman's  $\rho$ ,  $n=48$ ,  $r=0.63$ ,  $p<0.001$ ; Figure 5.2). The significant positive correlations indicate agreement between the observer's and the taster's scoring.

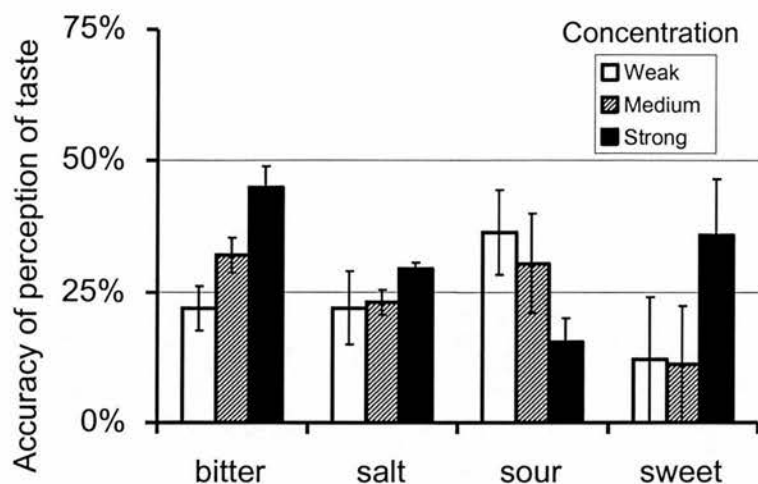
### **Taste strength**

There was a significant correlation between the observers' scoring of taste strength and the tasters' scoring of taste strength (Spearman's  $\rho$ ,  $n=48$ ,  $r=0.65$ ,  $p<0.001$ ). The correlation between the tasters' scoring of taste strength and the actual taste concentration was significant but weaker than might be expected (Spearman's  $\rho$ ,  $n=48$ ,  $r=0.29$ ,  $p=0.044$ ). There was no significant correlation between observers' scoring of taste strength and the concentrations of the solutions (Spearman's  $\rho$ ,  $n=48$ ,  $r=-0.06$ ,  $p=0.95$ ).

### **Taste judgements**

75% of the time tasters were correct in their judgements of the taste. Observers were however, poor at guessing which type of taste the drinks had from the tasters' behaviour. For each trial observer's correct guesses were given a value of 1 and incorrect guesses a value of 0. The proportion of observations correct for each of the 48 tasters' trials was then calculated. Observers estimated the actual taste correctly 29.5% of the time. This is significantly above chance (t-test against a hypothesised mean of 25%;  $t_{41}=2.19$ ,  $p=0.034$ ). The observers also estimated the taste that the taster judged the drink to have with an accuracy significantly above chance (29.9% of the time;  $t_{41}=2.31$ ,  $p=0.026$ ).

Further analysis revealed that this accuracy was mainly due to observers estimating significantly above chance for the sweet ( $t_{11}=2.94$ ,  $p=0.013$ ) and bitter stimuli ( $t_{11}=2.29$ ,  $p=0.043$ ) but not for either sour ( $t_{11}=0.213$ ,  $p=0.835$ ) or salt stimuli ( $t_{11}=-0.11$ ,  $p=0.913$ ). Accuracy scores may not reflect subject sensitivity to the presence of different stimuli since they do not take into account the number of false



**Figure 5.3 Accuracy of observers' perception of taste. Average accuracy (% correct  $\pm$  1 S.E) across 26 observers classifying tastes from video sequences of 4 subjects tasting liquids of different weak, medium and strong concentration.**

positive errors that subjects make. To take into account effect of false positives,  $A'$  was calculated for observers recognition of each taste (Snodgrass, Levy-Berger, & Haydon, 1985). The  $A'$  results are similar to the results of the analysis of accuracy above: salt and sour both have low values 0.27 and 0.26 respectively and sweet and bitter having scores of 0.70 and 0.56 respectively.

The accuracy of taste judgements was also affected by the solution concentration as can be seen in Figure 5.3. There were no significant correlations between solution concentration and the proportion of correct taste estimates for sweet (Spearman's rho,  $n=12$ ,  $r=0.00$ ,  $p<1.00$ ), sour ( $n=12$ ,  $r=-0.13$ ,  $p=0.68$ ) or salt stimuli ( $n=12$ ,  $r=0.30$ ,  $p=0.341$ ) but there was a significant correlation for bitter stimuli ( $n=12$ ,  $r=0.88$ ,  $p<0.001$ ). This correlation does not appear to be due to observers using hedonic value to guide judgements since when the correlation is repeated partialling out observers ratings of hedonic value, the result remains significant ( $r=0.79$ ,  $df=9$ ,  $p=0.004$ ).

## Discussion

To summarize the results: observers were able to infer how much the taster had enjoyed the drink from watching the video sequences; observers were also able to

make judgements of the strength of the drink but were generally poor at making judgments of the drink's taste.

It has been suggested that facial signals only communicate information related to the hedonic value of a flavour rather than anything specific about the flavour itself (see Berridge, 2000, for review). The results of the current study confirm the findings of Rosenstein and Oster (1988) in suggesting that this view may be slightly over simplistic and that observers may be able to detect face reactions made to bitter stimuli.

Communication of information related to bitterness may be important since, as Galef (1995) noted, foods that humans described as sweet tend to be consumed preferentially by other animals whereas foods described as bitter by humans tend to be rejected by other animals. So being able to detect whether a conspecific was eating food that seemed bitter may give an observer cues to the palatability of a food source without the observer directly ingesting the food.

While consuming food many animals make stereotypical behaviours to liked foods. Pelchat *et al.* (1983) examined the facial reactions of rats to sugar solution paired with one of two conditions: an aversive stimulus, electrical shock, or with a stimulus that probably caused the rat to feel nauseous, Lithium Chloride. They found that the rats had different facial and bodily reactions to the two pairings. Electrically shocked rats made positive 'ingestive' facial reactions to the sugar solution whereas the nauseous pairing made the rats 'gape' (a rat facial disgust response) and avoid the food in other conditions. Whether other rats make use of these facial responses is debatable since experimental manipulations have not separated facial reactions from other possible means of communication (e.g. Hishimura, 2000).

Even if it is possible for an individual to read another conspecifics' behaviour and tell whether a food is bitter or not, detecting bitterness does not seem to be a particularly good basis of a method for avoiding toxic novel foods. The movement of a species that eats a wide range of foods to an unfamiliar ecosystem often results in poisoning through the consumption of novel toxic plants (Galef, 1995). So detecting bitterness is probably not a very good means of judging whether a novel food is toxic or not. It is possible that perception of bitter taste may instead have a role in the assessment of whether known foods are currently at their best.

The presence of bitter tasting compounds is not simply related to the food 'going off'. Complex organic molecules (e.g. compounds with phenolic groups) accumulate in plant tissue over time. Such compounds may be by-products of essential biochemical processes, or may be synthesized to discourage animal consumption because such compounds have a bitter taste and toxic effects. Whatever the reason, new growth often has the lowest concentrations of such secondary compounds. For diets that include leaf, fruit, shoot or root components, bitterness can therefore indicate the level of toxins. Bitterness is only part of a 'guide to food' equation since new growth may also have least nutritional reward. Deciding which plant parts to consume involves a cost benefit analysis of the level of digestible carbohydrate, protein and fat resources gleaned from consumption, and the level of toxins to be suffered. Does this food taste less or more bitter than would be expected for an ideal state of ripeness for food reward? Acquisition of such information may be facilitated through social learning including the recognition of the bitter reactions of experienced consumers.

So bitterness may be useful an index of a particular foods benefits even when there is no danger of being poisoned. Allocation of time budgets for foraging,



processing and the digestion of food items is an optimisation problem that is critical for all animals. Time can be wasted acquiring low nutritional value items and time can be lost while such items pass through the digestive track. Skilled foraging again can be guided by attention to taste reactions of others. Even the absence of reactions may indicate that an expected bitter taste is absent and the state of ripeness is optimal.

The present study's results suggest that bitter taste reactions may be communicated (shown in the actions of tasters and recognised by observers). Although others' reactions to bitter tastes may be recognisable, our study shows that we are better at recognising the taster's reaction to the hedonic value associated with food than their reactions related to food bitterness.

Other animals apart from humans also use of visual cues from the behaviour of conspecifics to assess the attractiveness of a food source. This type of observational learning of whether a food is good to eat has been documented in various species. Mason *et al.* (1984; 1982) found that two species of blackbirds (*Agelaius phoeniceus* and *Quiscalus quiscula*) could acquire food avoidance from observing an individual of their own or the other species eating the food and becoming ill. Fryday and Greigsmith (1994) found that house sparrows' (*Passer domesticus*) food choice was influenced not only by observing another sparrow eating quinine-treated food but also by seeing the behaviour of another sparrow that had been previously exposed to quinine-treated food encountering the same food again even if it was now unadulterated.

It might be expected that non-human primates would also use visual signals to gain information about food value from conspecifics and that the ethological literature would have many examples of such behaviour. Only recently, however, has food value communication been examined and only one research group has reported non-

human primates using observational learning from conspecifics in order to judge food value. Snowdon and Bow (2000) found that cotton-top tamarins (*Saguinus oedipus*) communicate information about food palatability in a laboratory setting. So when one of their preferred foods (tuna fish) was adulterated with pepper only 25% of the group ever actually tried it. According to the authors, communication was primarily through visual signals of disgust. Similar mechanisms are almost undoubtedly present in humans and play important roles in choosing nutritious food and avoiding disease in a world that is more dangerous than we might think. For example, Marks *et al.* (2000) found that one vomiting person at the restaurant could infect people across a whole restaurant so well-developed disgust recognition can be highly functional.

**Chapter 6: A test of perception of neutral expressions:  
normal and depressed subject performance.**

## **Introduction**

Facial expressions are a highly important means of communicating social information in humans. In this chapter a new method for examining the perception of expression is described and used to compare the perceptions of what is a neutral expression in normal and depressed subjects. Previous literature has shown depressed and non-depressed individuals differ in their perception of expressions. These differences may be very broadly split into three types: findings that depressed individuals are generally worse at perceiving facial expressions (Cooley & Nowicki, 1989; Gessler, Cutting, Frith, & Weinman, 1989; Persad & Polivy, 1993; Walker, McGuire, & Bettes, 1984); findings that depressed individuals see emotionally expressive faces as looking more sad (George et al., 1998; Gur et al., 1992) and findings that depressed individuals see emotionally expressive faces as more negative (Bouhuys, Greet, & Mersch, 1997; Hale, 1998). This literature will be briefly reviewed before a new methodology is introduced aimed at differentiating between these findings.

A number of papers investigating expression perception and recognition (Cooley & Nowicki, 1989; Gessler et al., 1989; Persad & Polivy, 1993; Walker et al., 1984) have found that depressed subjects are worse at appreciating facial expressions in general (rather than having problems with specific expressions). Walker *et al.* (1984) compared normal, schizophrenic and depressed subjects and found that depressed subjects were less accurate than normal subjects: in discriminating whether two presented expressions matched; in labelling presented expressions and in selecting from a bank of images the face corresponding to an expression word label. Depressed subjects were also found to be poorer at a facial identity discrimination task.

Depressed subjects have been shown to be slower in their response to facial expressions. Cooley and Nowicki (1989) compared non-clinically depressed and non-depressed subjects' performance on two discrimination tasks. In one task, subjects had to decide whether pairs of emotional faces were the same or different and, in a second task, whether pairs of words were the same or different. They found that depressed subjects were slower than controls at facial expression matching decisions but as fast as controls when deciding whether pairs of words were similar or not. Cooley and Nowicki (1989) concluded that depressed individuals might have difficulty in social interactions because they cannot 'keep up' with the numerous non-verbal facial cues that are generated and that this could be a factor in maintaining their depression.

From studies using mood induction there is evidence that depression relates to a more specific change in the perception of expressions. Bouhuys *et al.* (1995) used music to induce happy or sad mood states in subjects. Subjects were presented with 12 line-drawn schematic face chimaeras whose features portrayed different expressions through the shape of the mouth (a single flat line, an upward or downward facing semicircle) and the position and orientation of the straight line portraying each of the eyebrows. Subjects were asked to rate each face for the amount of fear, happiness, anger, sadness, disgust, rejection and invitation present. Bouhuys *et al.* found that subjects who were induced to feel sad rated less intense, ambiguous facial expression chimaeras as being more sad and rejecting, and rated the less intense expressions as being less inviting and happy compared to subjects induced to feel happy. Another study using mood induction also found emotionally congruent effects (i.e. the tendency for individuals to see the emotion that they are feeling in others' faces). Niedenthal *et al.* (2000) used music to induce happy or sad moods in subjects. Subjects were then presented with an expression continuum, which either started with

a face looking sad that changed in steps to a neutral expression, or started with a happy expression that changed to a neutral expression. Subjects were asked to stop at the step in the continua when the facial expression first changed from that of the starting face. Niedenthal *et al.* found that emotionally congruent expressions persist longer than emotionally incongruent expressions. It may be that the effect of mood induction seen above is to shift what subjects see as being neutral, so a face seen as being neutral by happy subjects is seen as sad by sad subjects.

Depressed subjects also show mood congruent effects and rate faces as being more sad than non-depressed subjects (George *et al.*, 1998; Gur *et al.*, 1992). Gur *et al.* (1992) found that depressed subjects rate happy, neutral and sad faces as being less happy / more sad than normal subjects when rating faces on a scale from happy to neutral or sad to neutral. George *et al.* (1998) found that a rapid cycling, manic-depressive patient judged neutral and sad faces as being more sad during depressed phases than during non-depressed phases. It should be noted that in both of the above studies subjects rated facial expressions on only one dimension relating to sadness and/or happiness.

Other studies comparing the facial expression perception of depressed and control individuals have found that depressed individuals have a bias to rate expressions as more negative in general rather than just more sad. These studies have included many by Bouhuys *et al.* (1999a; 1999b; 1996; 1997). These studies found that clinically depressed subjects tended to rate line-drawn schematic face chimaeras (as described above) higher on a general negative scale than normal subjects. The negative scale was calculated as the average of subjects' ratings of face stimuli on fear, anger, sadness and disgust. Bouhuys *et al.* found that a tendency to make negative attributions to schematic faces seemed to correlate positively with the



probability of continued depression and relapse of depression in female subjects. In contrast, the tendency of male subjects to make negative attributions to schematic faces seemed to correlate negatively with continued depression (Bouhuys et al., 1999b). Hale (1998) found that depressed subjects rate the same line-drawn, schematic face chimaeras as being more negative (higher ratings on sadness, anger, disgust, fear and rejection) than control subjects. Hale also found that partners of individuals suffering major depression rate schematic, line-drawn face stimuli as less positive than controls.

The effects shown in the studies above could have common roots. Findings that depressed subjects are worse at facial expression tasks may in part be due to depressed subjects seeing faces as being sadder or more negative and therefore scoring more errors overall. In order to clarify the position three testable hypotheses were proposed: (1) depression leads to generally poor perception of facial expressions; (2) depressed individuals see faces as looking generally more sad; and (3) depressed individuals see faces as generally more negative (more angry, sad, disgusted and possibly more fearful).

In order to differentiate between these hypotheses a new test was created in which subjects are asked to make an emotional face image, presented on-screen by computer, look neutral. The subject can change the image by moving the mouse from left to right. This movement changes the on-screen image using image morphing. The change in the individual image is related to adding or diminishing the amount of anger, disgust, fear, happiness, sadness or surprise in the on-screen face image. The amount of anger, disgust, fear, happiness, sadness or surprise chosen by the subject when they have made the on-screen image as neutral as possible was recorded.

Each of the 3 hypotheses predict different results from the new test. The first hypothesis suggests depressed subjects have an overall deficit in the ability to deal with expressions, rather than a bias in either perception or labelling. If this hypothesis were correct it would be expected that depressed subjects would be poorer at seeing the emotional expression in the faces so, when presented with a series of trials in which they were asked to make a face look neutral, they would be less consistent in where they selected the emotionally neutral point (i.e. they would perform with greater variance between similar trials than non-depressed subjects). In order to test this, each subject performed the test in 2 blocks. In each block they were presented with the same trials allowing a subject's consistency to be studied across blocks.

The second hypothesis, that depressed subjects see faces as looking sadder than normal subjects, predicts that depressed subjects will choose less sad facial expressions as neutral, compared with normal subjects. This is because if depressed subjects tend to see facial expressions as being more sad then, in order to see a face as being neutral, they should need a face with less cues to sadness than normal subjects. One would expect the same effect if the third hypothesis is correct but depressed subjects should also choose faces showing less of other negative expressions (anger, disgust and possibly fear) than non-depressed subjects when selecting the most neutral face.

The test is based on Ekman and Friesen's (1976) standardized set of expression faces. Subjects use interactive computer graphics to choose the face that looks most neutral. Results of previous studies of facial expression perception may have been influenced by depressed individuals (and subjects with an induced, negative mood) being biased to use negative labels and to avoid positive labels. Such a bias can be seen in studies of subject affect and word recognition (e.g. Niedenthal &

Halberstadt, 1997). To avoid such influence, our task minimises the use of verbal labels and the expressions presented to subjects are made less visible, and hence less likely to be explicitly categorized by subjects, by decreasing their intensity. Neutral faces possess few cues to an individual's feelings, and are in that way ambiguous. Bouhuys *et al.* (1995) found the largest difference between depressed and normal subjects' expression ratings of faces occurred when using ambiguous face stimuli. Erroneous interpretation of ambiguous situations has also been suggested to be of causal importance in depression (Beck, Rush, Shaw, & Emery, 1979).

It was expected that at least some normal subjects would perceive the neutral faces of Ekman and Friesen (1976) as not being "neutral", possibly being more sad, angry, etc. than their own percept of neutrality. Subjects were therefore allowed to move along the expression continua beyond the original neutral point. To enable this a technique based on image morphing was used. Image morphing is used to interpolate a smooth linear change between two original images. To prepare a facial image for image morphing the x,y co-ordinate positions of corresponding salient features are recorded. The position of the features in a new image, say 75% of the way between the 2 original images, can then be calculated by adding 25% of the co-ordinate positions from the first image to 75% of the co-ordinate positions of the second image. The original images are then distorted using image warping into the new shape. Colouration of the final image is calculated as a fade 75% of the way from the first to the second distorted image. The texture of the resulting image is slightly blurred owing to elements of the textures not aligning. To solve this problem image texture processing based on wavelet magnitudes of the images (Tiddeman, Burt, & Perrett, 2001) was used. These techniques make a smooth linear transition between the two images. Since the transition between the images is linear it is possible to

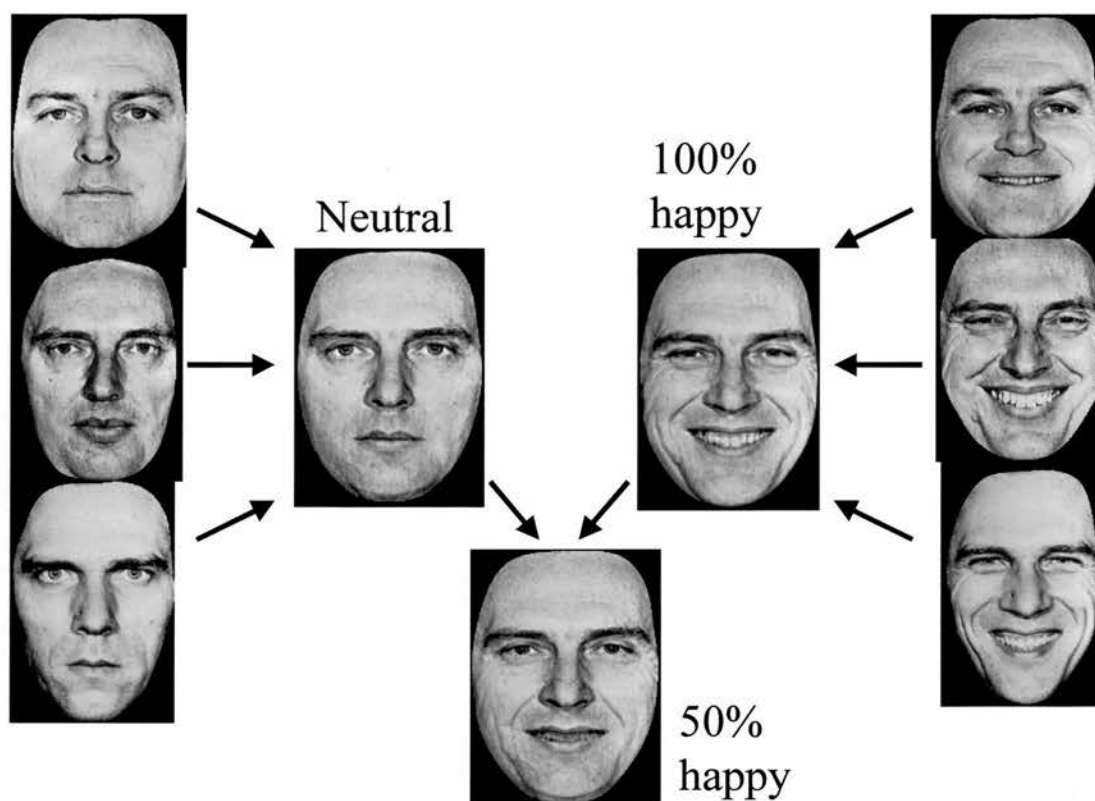
extrapolate the linear changes that happen between the images into a hypothesized region of face space outside the original image pair. This extrapolation technique is used to manufacture ‘anti-expressions’, images that form part of a linear morph continuum that starts with an expression which is morphed towards a neutral expression, but rather than stopping at neutral the changes are continued in the same direction to form an ‘anti-expression’. Hence in this expression to anti-expression continuum, subjects can find the facial configuration that looks most neutral.

## ***Methods***

### **Experimental set-up**

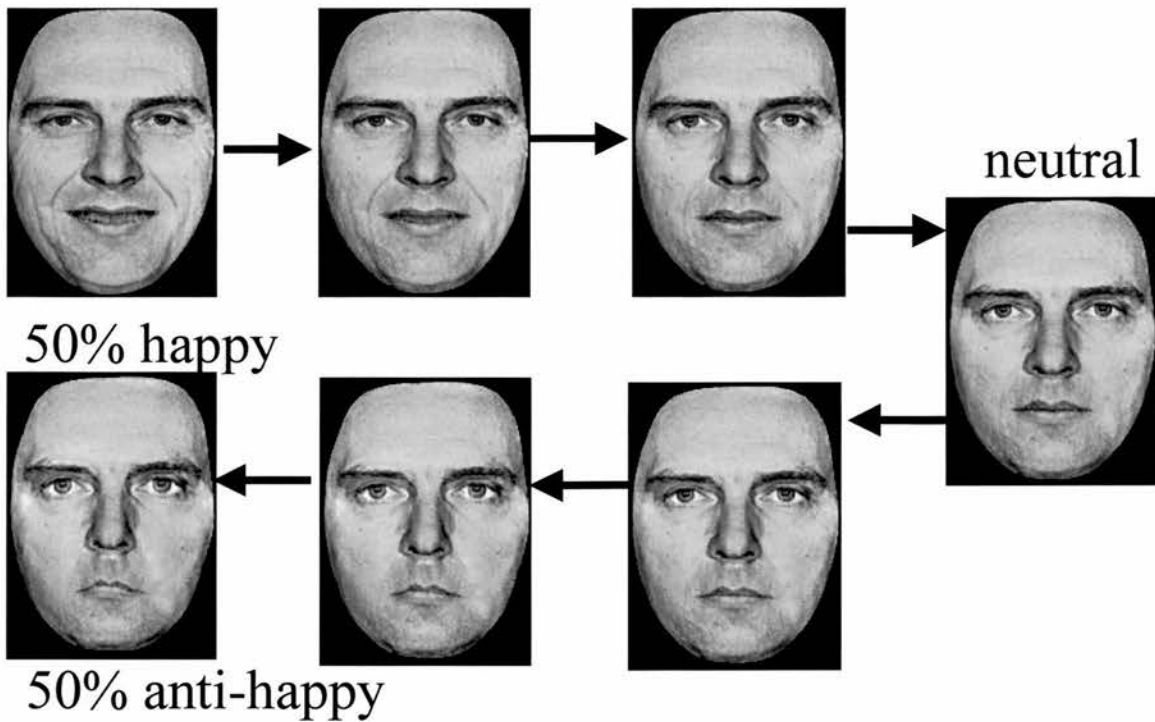
#### **Manufacture of image stimuli**

The images of nine people using the Facial Action Coding System to pose expressions of anger, disgust, fear, happiness, sadness and surprise from Ekman and Friesen (1976) were delineated. These 9 individuals’ images were used to make three composite identities: female 1 (individuals with initials: C, MF, and MO), female 2 (individuals with initials: NR, PE and PF) and the male composite (individuals with initials: EM, JJ and WF). For each of the composite identities 7 images (6 expressions and neutral) were made by blending 3 individuals’ images with the same expression. Blending each expression image with the corresponding neutral image was used to halve the intensity of expression in each of the composite images. This process is illustrated in Figure 6.1. This resulted in 18 images depicting expressions at +50% intensity and 3 neutral images, which were then used to generate the test image sequences. Composite facial expressions were used to minimise artefacts produced by the image manipulations.



**Figure 6.1** Manufacture of male composite individual neutral and 50% images. 3 neutral face images (far left) were combined to make a new male neutral face image. In the same way 3 happy face images (far right) were combined to make the a new 100% happy image. The 100% happy image and the neutral image were then combined to make a +50% happy image. This and the neutral image were then used to make the stimulus image sequence.

For each of the 3 composite identities, 6 image sequences were generated for each expression. The process is illustrated for the sequence involving the male composite happy expression in Figure 6.2. For the expression, the +50% intensity blend and the corresponding neutral face were used to make a linear sequence of 31 frames. These ranged from frame 1, the 50% expression stimulus through 14 intermediate ‘morph’ steps to the corresponding neutral image (frame 15). The sequence was then extrapolated through 15 frames to a -50% ‘anti-expression’ image.



**Figure 6.2** Manufacture of male happy stimulus image sequence. Using the male +50% happy and the neutral images a series of images were made using image morphing ranging from the +50% happy through to neutral. Caricaturing was then used to continue the shape, colour and texture changes to 50%, anti-happy. In the experimental sequence there are 31 frames, only 7 are illustrated here.

All images were masked to remove parts outside the face including the hair and neck, then resized to approximately 300 by 420 pixels.

Three practice sequences were also manufactured one for each of the composite identities. Each of the practice sequences started with a mix of the +50% surprised and +50% fear face and contained a linear sequence of 31 images that went through the neutral face image to -50% anti-(surprise-fear).

### **Stimulus presentation**

Subjects were presented with a set of instructions and then with 3 practice trials using the same interface that would later present the actual stimuli. The instructions told the subjects that they were to see the faces of 3 different individuals making slightly different expressions and that they were to move the mouse from left



to right to make the face of each individual as neutral as possible. They were told that neutral meant not looking happy, sad, surprised, angry or afraid, but just having a relaxed expression, and that often the facial expressions will look very similar. After the practice trials, subjects went onto the first, followed by the second, of 2 blocks of trials in which all 18 continua were presented in random order. Each trial consisted of a single continuum, subjects were asked to move the computer mouse horizontally until they found the most neutral looking face. Subjects were instructed to click with the mouse once they had found the most neutral looking face. This caused the computer program to record the current position in the sequence and display the next trial.

During a trial horizontal mouse movement produced a real-time change in the face image presented. Mouse movement from one side of the screen to the other produced a stepped increase in expression from  $-50\%$  to  $+50$  followed by a stepped decrease to  $-50\%$ . To prevent subjects from using the position of the mouse relative to the screen centre as a guide to the location of the neutral point in the image sequence, the relationship between mouse position and image presented was varied randomly between trials. This was done by offsetting the start of the image sequence by the random amount (% of the screen) and wrapping the end of the sequence round to the start of the screen.

## **Experimental testing**

### **Participants**

34 female and 25 male subjects with average age 44 years (age range 22 to 73 years) from general populations of Scotland and Italy.

## **Selection of Low and high BDI groups**

To distinguish between depressed and non-depressed subjects we chose to use Beck's Depression Inventory (BDI, Beck, Ward, Mendelson, Mock, & Erbaugh, 1961). The BDI originated as a tool for use for pre-diagnostic screening, and also for the quantification of level of depression in individuals already diagnosed with depression. The BDI is often used to distinguish a depressed sample of individuals from a non-diagnosed group of individuals for experimental purposes. There is debate about the comparability of clinically diagnosed, depressed individuals and individuals that have been selected as being depressed on the basis of the BDI (Cox, Enns, & Larsen, 2001; Gotlib, 1984). Gotlib highlights a problem of the BDI scale being a self-report measure in that many studies of non-clinical groups have found BDI score to be related to other self-report measures, particularly those measuring anxiety. Gotlib also notes, however, that the inter-correlation of BDI score and anxiety has been reported for samples of psychiatric patients.

Recently, Cox *et al.* (2001) compared the self-reported, symptom severity profiles of patients with a major depressive disorder and student subjects who scored high (above 16 or above 21) on BDI. Cox *et al.* found that student subjects with high BDI scores were similar in symptom severity to the clinical group and recommended the use of higher BDI cut-offs to classify depression than the currently most often used score of 10. Beck and Steer (1987) also advised the use of high BDI cut-offs to detect possible depression. We therefore chose to use a BDI score of 16 as the cut-off point between groups of 'Low' BDI (non-depressed) and 'High' BDI (depressed) subjects.

## **Procedure**

Subjects read and signed a consent form, after which they completed Beck's Depression Inventory, followed by the neutral expression experiment.

## **Consistency**

Subjects saw each trial twice, once in the first block and once in the second. If subjects were highly consistent in their choice of neutral face, their choice would be similar in the first and second blocks. On the other hand, if the subject is inconsistent then there would be large variation between like trials in the two blocks. We calculated 'Consistency' for each expression as the average difference across the three identities between same stimulus trials in the first and second blocks of the experiment. Overall Consistency was calculated as the average Consistency score for all of the expressions.

The subjects had an average Overall Consistency of 0.15, sd 0.039 (high numbers indicate less consistency). One of the subjects (cm24) showed much lower Overall Consistency than any of the other subjects. This subject had a value of 0.31, over 4 standard deviations away from the mean and 2 standard deviations away from the next highest value. This low consistency subject, a male with Low BDI was eliminated from further analysis.

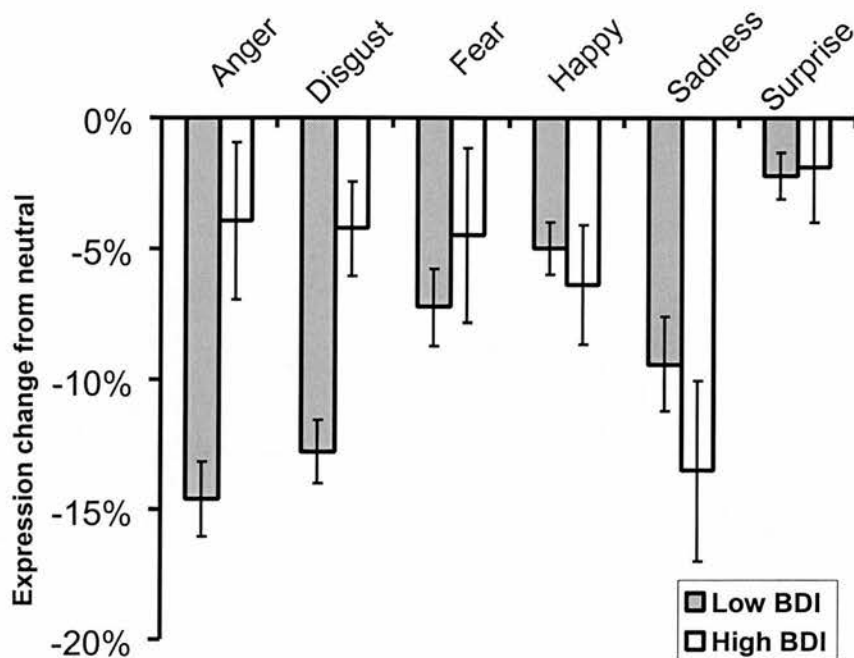


Figure 6.3 The average amount of expression chosen by control and depressed (Low and High BDI) subjects to make the faces look neutral, plotted with standard error bars.

## Results

### Overall differences from neutral

When asked to choose the most neutral face, subjects were biased to choose not the position of the original 'neutral' face from the continua (i.e. 0%), but selected faces transformed towards the anti-expression. This was true of the continua representing each expression: anger ( $t$ -test against a null hypothesis mean of 0% transform:  $t_{63}=-9.33$ , mean=-0.12, sd=0.097,  $p<0.001$ ), disgust ( $t_{63}=-9.91$ , mean=-0.11, sd=0.077,  $p<0.001$ ), fear ( $t_{63}=-5.15$ , mean=-0.06, sd=0.090,  $p<0.001$ ), happy ( $t_{63}=-5.68$ , mean=-0.05, sd=0.063,  $p<0.001$ ), sadness ( $t_{63}=-6.19$ , mean=-0.09, sd=0.112,  $p<0.001$ ) and surprise ( $t_{63}=-2.65$ , mean=-0.02, sd=0.060,  $p<0.05$ ).

### Differences between High and Low BDI subjects

A BDI score of 16 was used as the cut-off between Low BDI ( $n=57$ , average BDI score=4.2, standard deviation of BDI score=3.7) and High BDI ( $n=7$ , average BDI score=20.4, standard deviation of BDI score=4.3) subject groups. A repeated measures ANOVA was performed for point selected as neutral with BDI score (2

levels: Low and High) and Expression (a within subjects factor with 6 levels: anger, disgust, fear, happiness, sadness and surprise). The results are presented in Figure 6.3. There was a significant interaction between Expression and BDI score ( $F_{5,58}=3.50$ ,  $p<0.01$ ) and a main effect of Expression ( $F_{5,58}=4.74$ ,  $p<0.001$ ). The main effect of Expression was due to subjects choosing different amounts of transform to make the facial expression look neutral depending upon the facial expression.

To investigate this interaction a series of Post-Hoc tests were performed to compare the amount of transform chosen to make the face look neutral for each expression by the Low and High BDI groups. There were significant differences between the amount of anger and disgust transform that Low and High BDI groups subtracted from the neutral face image to make it look neutral. The High BDI group subtracted out less anger and disgust expressions than the Low BDI group (LSD Post-Hoc test for anger  $p<0.001$ ; disgust  $p<0.001$ ). There were no significant differences between High and Low BDI scoring groups for fear, happy, sad or surprise continua (fear  $p=0.418$ ; happy  $p=0.444$ ; sad  $p=0.084$ ; surprise  $p=0.859$ ).

## **Consistency**

To test whether High and Low BDI scoring subjects differed in how consistent they were in their choice of neutral face, a repeated measures ANOVA was performed for Consistency with BDI score (2 levels: Low and High) and Expression (a within subject factor with 6 levels: anger, disgust, fear, happiness, sadness and surprise). There was a main effect of Expression ( $F_{5,58}=4.63$ ,  $p<0.001$ ) but no significant effect of BDI score ( $F_{1,62}=1.25$ ,  $p=0.266$ ) or interaction between BDI score and Expression ( $F_{5,58}=1.45$ ,  $p=0.220$ ). This indicates that there were no differences between High and Low BDI scoring individuals in terms of their Consistency.

## ***Discussion***

To make the faces look neutral, High BDI, 'depressed', subjects transformed on-screen facial images by subtracting less anger and disgust expressions out of the images than Low BDI subjects. It thus appears that in comparison to Low BDI subjects the High BDI scoring subjects saw more angry and disgusted faces as being neutral. The High BDI scoring subjects were as consistent as the Low BDI scoring subjects in their choice of where neutral was. These results do not fit with any of the hypotheses proposed at the outset from the facial expression and depression literature. Later in the discussion the findings will be related to other literature on the reactions of depressed individuals.

No evidence was found to support the first hypothesis that the High BDI subjects are poorer in expression perception, by being less consistent in selection of where neutral faces lie, than Low BDI subjects. Rather, the High BDI subjects seem to be as consistent in their selection of where neutral faces lie on the image continua as other subjects. It is in many ways unsurprising that very highly depressed subject would have poor performance on expression tasks, but some previous studies that have found that depressed subjects are worse at facial expression related tasks have used non-clinical subjects who are probably less depressed and more comparable to the subjects used here. Instead it must be concluded that either our measure of consistency is not sensitive enough to differentiate between Low and High BDI subjects or that differences in task demands led to the differences between the current and previous studies' results.

Overall, subjects chose to subtract expressions out of the face stimuli when asked to make faces look neutral, as can be seen from Figure 6.3. There may be two possible reasons for this. The neutral faces in the Ekman and Friesen (1976) set are



sometimes said to look slightly negative. This might be due to their expressions being neutral in muscular configuration rather than social meaning. In normal social interaction people may be slightly positive in their facial expressions. Thus, a technically neutral pose may be seen to be slightly anti-social. Although this interpretation can account for subjects subtracting out the negative expressions of anger, sadness and disgust, it cannot account for the smaller but still highly significant effect of subjects transforming the face to subtract out the happy expression. The bias of neutral point away from real expressions is possibly due to 'anti-expressions' being less expressive, because they are artificial and so are not encountered in the natural world. When subjects perform the experiment, they may search for the neutral expression by trying to avoid recognisable expressions. Since anti-expressions are less 'expressive' and have less emotional meaning than real expressions, subjects may be biased to finding neutral further from real expressions than anti-expressions.

To shed light on what was driving the above effect, a short test was performed in which subjects rated the ends +50% and -50% (anti-expression) of the continua and the neutral face for each expression (see the appendix at the end of the chapter) on expression intensity. The results show that for all of the expressions, apart from happiness (for which the anti-expression was rated as showing less happiness than the neutral), the neutral face was rated the same as the anti-expressions. This supports the proposition that subjects tend to select anti-expression faces as being neutral because they are avoiding the more expressive, positive parts of the emotional ranges.

The introduction presented 3 hypotheses. The first hypothesis was not supported by the data as already discussed. The second and third hypotheses held that depressed subjects would take out more sadness, or more negative expressions in general, from faces than non-depressed subjects when making face images look

neutral. Neither of these hypotheses were confirmed; instead our results may be interpreted as suggesting that depressed subjects' percept of what is a neutral facial expression is more angry and disgusted than that of non-depressed subjects. This result may relate to the different experiences of the world by high and low BDI subjects. High BDI subjects may be exposed to more hostile reactions than low BDI subjects and this may affect their percept of what a neutral facial expression is like.

Studies have found that depressed subjects arouse more negative reactions from others. For example, Coyne (1976), comparing depressed and non-depressed subjects' telephone conversations, found that the mood of individuals interacting with depressed subjects becomes more negative than that of individuals interacting with non-depressed subjects. This led him to conclude that depressed people might induce depression and hostility in others. Later Gotlib and Robinson (1982) examined the interaction between female student pairs. Unlike Coyne, they found no effect of depressed subjects on the mood of the individuals with whom they were talking but that interactions with depressed subjects triggered more negative behaviour than interactions with non-depressed subjects.

Depressed subjects appear also to expect more negative expressions in different situations. Berenbaum (1992) asked depressed and non-depressed subjects to pose the expressions that they would use in imagined situations. He found that depressed subjects showed more anger and contempt and, when smiling, looked less genuinely happy than the non-depressed subjects. This behaviour of depressed subjects in imagined situations may not only indicate their normal reactions but also the behaviour they expect from others. If the negative reactions of others are the norm in depressed subjects' experience, then what depressed subjects expect as being neutral should be biased to reflect this. Thus, faces that depressed subjects see as

being neutral are more angry and disgusted than faces seen as neutral by control subjects.

Many previous studies have asked subjects to label or rate emotions in facial images and found that depressed subjects thought that facial expressions were more sad (George et al., 1998; Gur et al., 1992) or more negative (Bouhuys et al., 1999a; Bouhuys et al., 1999b; Bouhuys et al., 1997; Hale, 1998) than non-depressed subjects. This may be a result of subjects projecting their own emotions onto the faces pictured. The present study asked a different question, “What do subjects perceive as being neutral?”.

The results of the experiment are interpreted as reflecting the facial expressions that are normally encountered by depressed and control subjects. Depressed subjects may be more likely to encounter or trigger negative reactions from others; they may therefore treat slightly negative expressions as normal. The test may demonstrate that a depressed individual’s concept of what is a neutral facial expression is more angry and disgusted than that of other individuals.

The new methodology proposed here may be used to establish whether hostile interactions are present in the environment of depressed individuals who have become less depressed. Previous research has suggested that persistent hostile interactions between individuals and their environment may be a cause of depression (Lewinsohn, 1975).

## ***Conclusions***

This chapter presents a novel test method that may be used to investigate the differences between the perceptions of what is a neutral facial expression in different groups of subjects. The method’s application is illustrated in the investigation of depressed subjects by showing that their perception of what constitutes a neutral face

is more disgusted and angry than that of non-depressed subjects. These results differ from the previous literature on facial expression. The finding is interpreted as reflecting differences in environment between depressed and non-depressed individuals. Depressed individuals may encounter more hostile reactions from others, leading to their percept of a neutral, normal facial expression being biased towards hostility.

## **Appendix**

To understand more about the anti-expressions a short test was set up to compare, from each continuum, subjects' ratings of the +50%, neutral and -50% frames. The test was performed by 8 female and 5 male student subjects (average age 25) who rated the 54 stimulus images on a 7-point scale anchored at each end by 'very' and 'not at all' and the emotion relevant to the continua: angry, disgusted, fearful, happy, sad or surprised. The stimuli consisted of 6 different expressions at 3 levels of intensity (+50%, neutral and -50%) posed by the same 3 composite individuals as used in the main experiment. The neutral and -50% stimuli did not differ in the amount of expression they were rated as expressing for anger (Wilcoxon signed ranks test comparing ratings for neutral and anti-expressions;  $N=13$ ,  $Z=-1.25$ ,  $p=0.21$ ), disgust ( $N=13$ ,  $Z=-1.31$ ,  $p=0.19$ ), fear ( $N=13$ ,  $Z=-0.10$ ,  $p=0.92$ ), sad ( $N=13$ ,  $Z=-1.10$ ,  $p=0.27$ ), or surprise ( $N=13$ ,  $Z=-0.36$ ,  $p=0.72$ ), but differed for expressions of happiness for which the -50% was rated as less happy than the neutral expression ( $N=13$ ,  $Z=-3.07$ ,  $p=0.002$ ). For all expressions the +50% stimulus face image was seen as having more of the expression than the neutral image and was seen as having more of the expression than the -50% image (for all comparisons;  $N=13$ ,  $Z<-3.00$ ,  $p<0.002$ ).

## **Chapter 7: Effects of adaptation to anti-face expressions on perception.**



## **Introduction**

Webster and MacLin (1999) found that adaptation by repeated viewing of a distorted face image caused subjects to judge subsequently viewed face images as most normal when they were distorted slightly in the direction of the adapting distortion. Following this work, Leopold, O'Toole, Vetter and Blanz (2001) investigated adaptation to 3D 'anti-face' stimuli manufactured to possess the opposite characteristics to a corresponding individual (Blanz, O'Toole, Vetter, & Wild, 2000). For example, if a particular individual's face has a large nose and dark eyebrows in comparison to average then the corresponding anti-face will have a small nose and light eyebrows. Leopold *et al.* (2001) trained categorisation of the face images of 4 individuals. During the test phase, subjects adapted to an anti-face stimulus for 5 seconds, and then categorised the identity of one of the 4 trained faces shown for 200ms. Adaptation to the corresponding individual's face increased recognition accuracy, but decreased the recognition accuracy for the other faces. For example, adaptation to 'anti-Adam' enhanced recognition of Adam's face, but led to worse recognition of George's face. The results are interpreted by Leopold *et al.* as caused by an after-effect due to adaptation of cells probably located in the Superior Temporal Sulcus (STS) to the anti-face. Studies involving single cell recording have suggested that independent groups of STS neurons may be sensitive to facial identity and expression (Hasselmo, Rolls, & Baylis, 1989; Perrett, et al., 1984). So, here the effect of adaptation to anti-face expressions on the recognition of emotional expressions is investigated.

Anti-face expression stimuli were manufactured by extending previously described morphing techniques (Benson & Perrett, 1991; Tiddeman, Burt, & Perrett, 2001). The difference between a full-blown (100% expression intensity) facial

expression (composites for anger, sadness, disgust and fear based on images from Ekman & Friesen, 1976) and a neutral (0%) expression was used to transform a neutral expression to create less intense versions of the expression and anti-expression (-50% intensity) images with the opposite characteristics to the original expression (Figure 7.1). The effect of adaptation to the different anti-expressions on recognition of expressions was examined in an experiment with 2 consecutively presented parts.

Part 1 of the experiment serves to test whether adaptation to anti-expressions leads to better recognition of subsequently presented corresponding expressions. The identity of the individual portrayed in the stimuli was varied to shed light on the order in which individual identity and expression are processed. In Bruce and Young's (1986) model of facial recognition, expression analysis is processed separately from identity. So, adaptation to an anti-expression 'posed' by one individual might be expected to enhance recognition of the corresponding facial expression when presented by another individual. Alternatively, expression identification may be processed after individual identity (Schweinberger, Burton, & Kelly, 1999), enabling individuals differences in resting expression to be accounted for during expression decoding, and thus, adaptation effect would be expected to be specific to the individual pictured in the face stimulus.

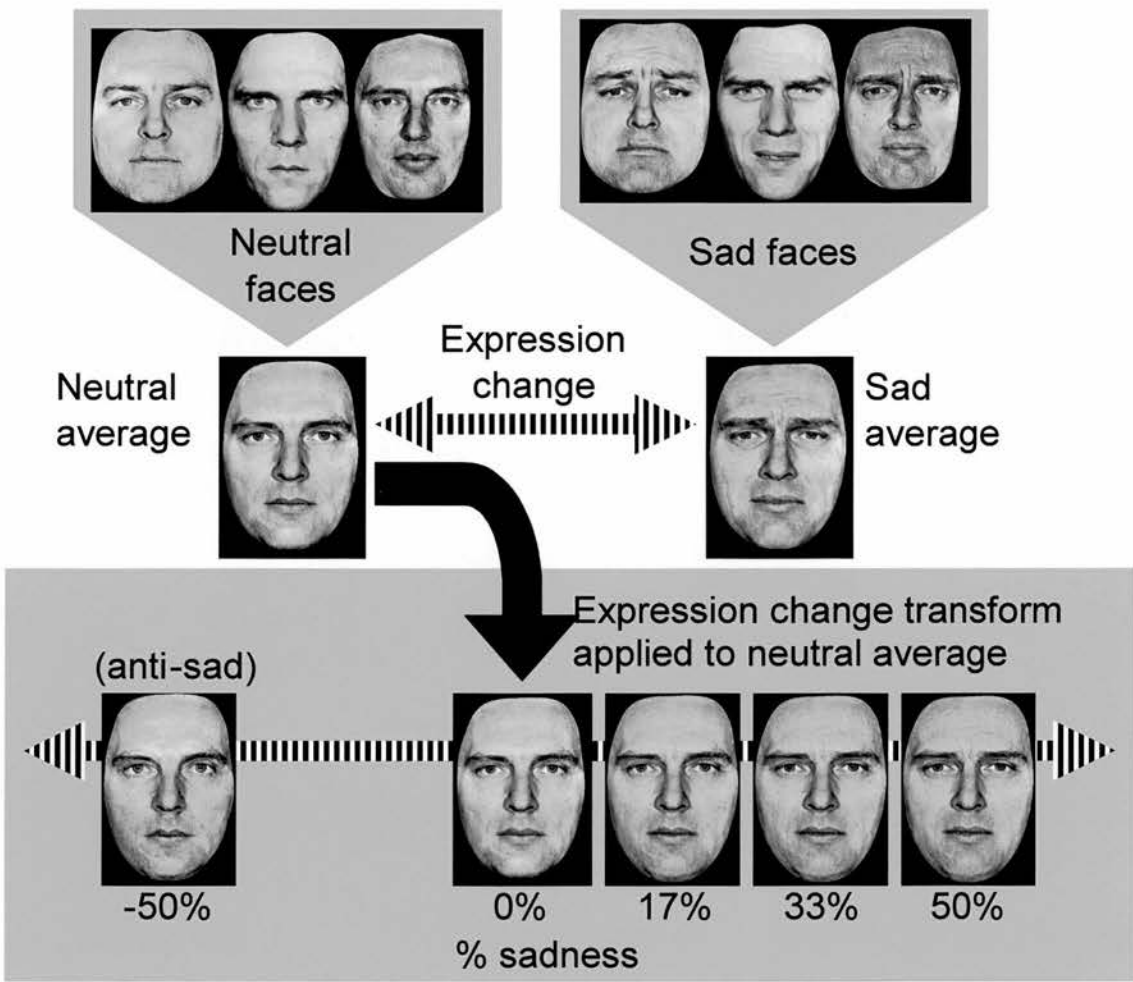
Bruce and Young (1986) suggest that studies in which the same face images are seen during both learning and test phases of recognition tasks may tap more general visual, rather than face specific, processing. This may explain why the highly trained subjects recognise the identity of face image stimuli so easily when the adapting and test faces were both rotated by 180 degrees (Leopold, O'Toole, Vetter, & Blanz, 2001). In the experiment presented here subjects were not trained to recognise

the face stimuli, as they are familiar facial expressions (Ekman & Friesen, 1976), and fewer trials were used to avoid over learning.

Flicking between similar images can cause apparent movement. An alternative explanation for the results found by Leopold *et al.* (2001) may be that apparent movement caused by the change between an anti-individual and the corresponding individual serves as a cue to identity. Apparent motion may be eliminated by the use of masking presented here to halt the processing of visual information. In the second part of the experiment masks are presented between the anti-expression and the expression to be identified. It is hypothesised that if the better recognition of the face image corresponding to the anti-face is a low-level visual processing effect it may then be disrupted by backward masking which has been shown to interrupt processing in STS (Keysers & Perrett, 2002; Kovács, Vogels, & Orban, 1995; Rolls, Toveé, & Panzeri, 1999). The effects of masking on the adaptation to anti-expressions will therefore be investigated using 3 different masks: a blank mask as a control, a noise mask to affect lower level visual processing and a more complex mask made from pieces of different faces to affect higher levels of visual processing.

In a supplementary experiment Leopold *et al.* (2001) investigated the effect of inducing a delay between the adapting anti-face identity and the test face and concluded that the facilitating effect of anti-individuals on the recognition of the corresponding individual faces is 'short-lived', being diminished after 300ms. In the second part of the experiment presented here, different levels of delay between the disappearance of the anti-expression image and the expression image are also investigated. It is hypothesised that these periods of delay between the offset of the anti-expression and the onset of the expression will decrease the facilitating effect of

accommodation to the anti-expression image because of the short lived nature of any adaptation effect.



**Figure 7.1** Manufacture of anti-expressions and graded intensities of expression. Average neutral and sad faces are manufactured by blending equivalent expressions of 3 individuals. The difference between these faces is applied to the average neutral face as a transform in luminance, shape and texture to generate a continuum for sadness running from +50% through neutral (0%) to -50% (anti-sad).

**Methods**

**Subjects**

Sixteen individuals (11 female; average age = 34, sd = 12).

**Materials**

A technique based on shape, colouration and texture (Benson & Perrett, 1991; Tiddeman, Burt, & Perrett, 2001) averaging and transformation was used to make

male and female -50% anti-expressions and varying intensities of anger, sad, disgust and fear (see Figure 7.1 and 7.2) based on the images of Ekman & Friesen (1976).

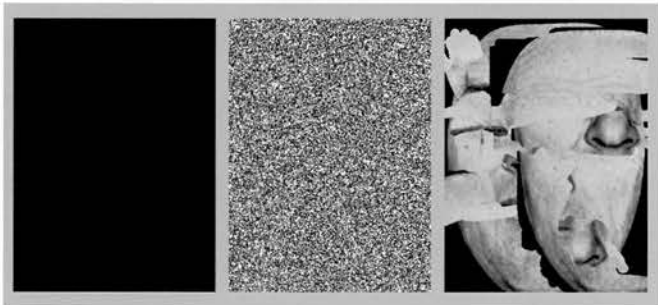
Software sequentially presented the anti-face for 2 seconds followed by the

test face for 500ms. The subjects' keyed response (keys A, S, D and F for Anger, Sad, Disgust and Fear respectively) led to a 3 second delay, after which the next randomly ordered trial was presented. (To maintain subject concentration and avoid eyestrain the software paused after every 10 trials until the spacebar was pressed.)

	-50%	+50%
Anger		
Disgust		
Fear		
Sadness		

### Practice

Subjects were asked to categorise the facial expressions of male stimulus faces that would appear briefly and were presented with 64 practice trials (16 of each of the 4 expressions at 50% intensity). During the practice trials, and Part 3, the anti-faces were not presented.



**Figure 7.2** Example anti-expression and 50% expression stimuli. For each of the 3 emotions example anti-expressions (-50%, centre) and +50% expressions (right) are shown. The masking stimuli are presented in the lower grey box from left to right: Blank, Noise (random noise) and Faceparts (made from jumbled selected face areas).

## **Part 1: Initial testing and robustness to change in face identity**

Subjects were told that they would see two face stimuli on each trial and were asked to look at the first face stimulus but to categorise the expression of the second face stimulus that appeared during each trial. Each subject performed 256 trials, selected from a total of 1024 possible trials, consisting of 2 identities of anti-expression stimulus \* 4 anti-expressions (anger, sadness, disgust, fear) \* 2 identities of expression stimulus \* 4 expressions (anger, sadness, disgust, fear) \* 16 levels of intensity (0%, 3%, 7%, 10%, 13%, 17%, 20%, 23%, 27%, 30%, 33%, 37%, 40%, 43%, 47%, 50%). Trials were split between 4 blocks and each subject only saw one expression at each level of expression intensity. Since subjects performed only a quarter of the trials multiples of four subjects were needed to counterbalance the experiment.

## **Part 2: Masking and delay**

Subjects were presented with instructions as in Part 1 followed by 192 trials in which the male face stimuli were presented. During each trial an anti-expression appeared followed by one of 3 masks (Figure 7.2) for a period of between 0 and 1 second and was followed by a facial expression<sup>24</sup>. The trials presented were as follows: 4 anti-expressions \* 4 levels of delay (0, 250, 500 and 1000ms) \* 3 mask types (Blank, Noise, Faceparts) \* 4 expression intensities (0%, 17%, 33% and 50%) \* 4 expressions. The trials were split and counterbalanced as above.

---

<sup>24</sup> During delay times of 0 seconds no mask was shown.



Part 3: Recognition of anti-expressions and debriefing

To test whether subjects could recognise the anti-expressions the subject practice phase was repeated with 64 trials presented for classification. These consisted of 2 expression intensities (50% and -50%, the anti-expression) \* 4 expression \* 8 repetitions of the male stimulus. After completion, subjects were questioned to ascertain their understanding of the anti-expression stimuli.

Results

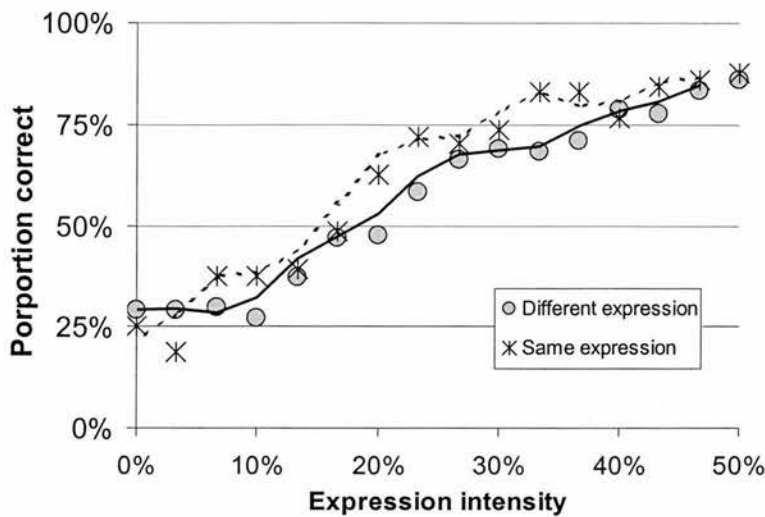
Overall, subjects were good at the practice task and correctly selected the appropriate expression for the image 77% on average of the time.

Part 1: Initial testing and robustness to change in face identity

A repeated measures ANOVA was performed on the proportion of expressions recognised correctly with *expression change* (anti-face and face same or different) as a within subjects factor and *individual change* (same or different individual pictured with the anti-face and face) and *intensity* as between subjects factors (see Table 7.1 for descriptive statistics). There was a significant effect of *intensity* ( $F_{15,480} = 49.0$ ,  $p<0.001$ ) caused by higher intensities of expression being more easily recognised and a significant effect of *expression change* ( $F_{1,480} = 8.37$ ,  $p<0.01$ ); reflecting the higher proportion of correct responses in trials in which the anti-face and the face represented

		Different individual		Same individual	
		Different expression	Same expression	Different expression	Same expression
Expression intensity	50%	84% (18%)	88% (29%)	88% (14%)	88% (22%)
	47%	84% (14%)	78% (31%)	82% (17%)	94% (17%)
	43%	78% (22%)	81% (25%)	77% (20%)	88% (22%)
	40%	77% (22%)	75% (41%)	80% (15%)	78% (31%)
	37%	71% (19%)	78% (26%)	71% (23%)	88% (22%)
	33%	75% (19%)	88% (29%)	61% (26%)	78% (31%)
	30%	67% (20%)	69% (31%)	71% (22%)	78% (26%)
	27%	59% (17%)	63% (34%)	73% (18%)	78% (26%)
	23%	57% (21%)	75% (26%)	59% (25%)	69% (31%)
	20%	51% (23%)	53% (46%)	44% (25%)	72% (31%)
	17%	45% (23%)	47% (43%)	49% (28%)	50% (37%)
	13%	35% (19%)	41% (38%)	39% (13%)	38% (34%)
	10%	28% (17%)	34% (35%)	26% (21%)	41% (38%)
	7%	31% (15%)	38% (39%)	28% (21%)	38% (34%)
	3%	31% (18%)	16% (30%)	27% (20%)	22% (31%)
	0%	30% (14%)	31% (31%)	28% (21%)	19% (25%)

Table 7.1 Average (and standard deviation) for each of the conditions in Part 1.



**Figure 7.3** Proportion of correct identifications in trials adaptation and test stimuli related to the same or different expression. To aid visualisation the 2 point moving average is plotted for same and different expression conditions. (The 2-point moving average is used only for this figure and not for the analysis.)

same compared to when they were different expressions (see Figure 7.3). There was no significant effect of *individual change* ( $F_{1,480} = 1.43, p=0.23$ ), although, the facilitating effect of the different individuals condition was 3%, compared to 7% when the individuals pictured were of the same identity. There were no other significant ( $p<0.05$ ) effects (all other  $p>0.20$ ).

## Part 2: Effects of masking and delay.

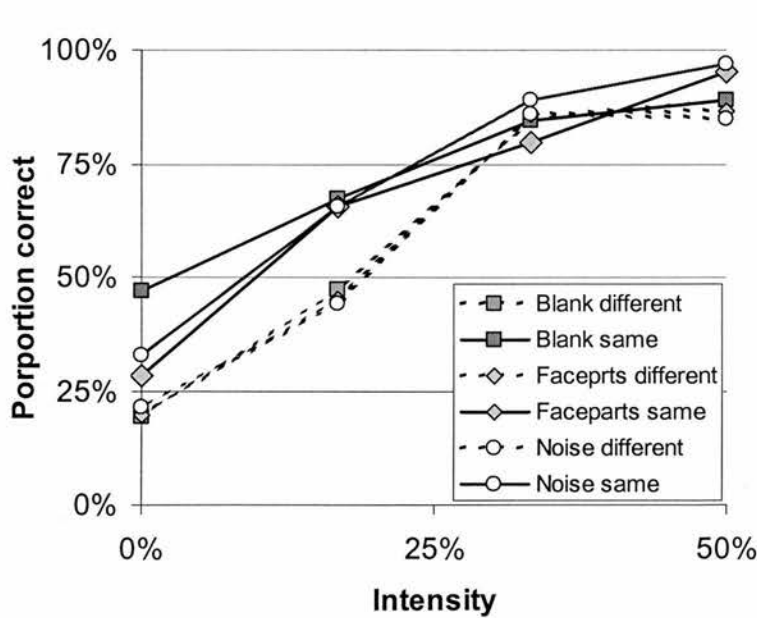
A repeated measures ANOVA was performed on the proportion of correctly recognised expressions with *expression change* (anti-face and face congruent or incongruent) as a within subjects factor, *mask type* (Blank, Noise and Faceparts), and *intensity* and *delay* as between subjects factors(see Table 7.2 for descriptive statistics).

Expression intensity	Delay (seconds)	Blank mask		Faceparts mask		Noise mask	
		Different expressions	Same expressions	Different expressions	Same expressions	Different expressions	Same expressions
50%	0	90% (16%)	88% (34%)	90% (16%)	94% (25%)	85% (17%)	94% (25%)
	0.25	88% (17%)	88% (34%)	83% (30%)	100% (00%)	94% (13%)	100% (00%)
	0.5	94% (13%)	94% (25%)	88% (17%)	88% (34%)	77% (32%)	100% (00%)
	1	85% (21%)	88% (34%)	85% (30%)	100% (00%)	83% (21%)	94% (25%)
	1	81% (21%)	88% (34%)	90% (16%)	81% (40%)	88% (21%)	88% (34%)
33%	0	81% (27%)	69% (48%)	83% (17%)	81% (40%)	90% (16%)	88% (34%)
	0.25	90% (20%)	94% (25%)	81% (24%)	81% (40%)	83% (27%)	94% (25%)
	0.5	88% (24%)	88% (34%)	90% (20%)	75% (45%)	83% (21%)	88% (34%)
	1	46% (30%)	75% (45%)	44% (29%)	63% (50%)	42% (31%)	81% (40%)
	1	52% (38%)	56% (51%)	50% (27%)	63% (50%)	44% (29%)	56% (51%)
17%	0	42% (31%)	63% (50%)	42% (31%)	69% (48%)	50% (27%)	56% (51%)
	0.25	50% (30%)	75% (45%)	46% (30%)	69% (48%)	42% (33%)	69% (48%)
	0.5	15% (21%)	50% (52%)	15% (27%)	31% (48%)	21% (21%)	31% (48%)
	1	21% (21%)	44% (51%)	19% (21%)	31% (48%)	17% (21%)	38% (50%)
	1	17% (17%)	63% (50%)	25% (23%)	38% (50%)	25% (19%)	19% (40%)
0%	0	25% (19%)	31% (48%)	21% (21%)	13% (34%)	23% (26%)	44% (51%)
	1						

**Table 7.2.** Average (and standard deviation) for each of the conditions in Part 2.

There was a significant effect of *expression change* ( $F_{1,240}= 21.60, p<0.001$ ) and of *intensity* ( $F_{3,240}= 161.44, p<0.001$ ), but no significant effects of *delay* ( $F_{3,240}= 0.09, p=0.966$ ) or *mask* ( $F_{2,239}= 1.40, p=0.25$ ). There was a significant interaction between *expression change* and *intensity* ( $F_{3,240}= 4.60, p<0.01$ ), possibly reflecting a ceiling effect at higher expression intensities, see Figure 7.4, and an interaction between *expression change*, *intensity* and *mask* ( $F_{6,480}= 2.33, p<0.05$ ) that is reflected in the interactions in the alternative analysis presented below. There were no other significant interactions.

An alternative analysis was performed splitting by *mask* condition, discarding trials in which *delay* = 0 (i.e. trials in which no mask was shown) and pooling across *delay*=250, 500 and 1000. For each mask condition a repeated measures ANOVA was performed on the proportion of facial expressions recognised correctly with *expression change* (anti-expression and expression same or different) as a within subjects factor and *intensity* as between subjects factor. Table 7.3 shows that there was a significant effect of expression change for each type of mask. For all ANOVAs



**Figure 7.4** The relationship between the proportions of correctly identified facial expressions and the intensity of expression for same/different anti-face trails under different masking conditions.




Mask	Accuracy (% correct)		Facilitation	F=	df=	p=
	Expression Different	Expression Same				
	61%	71%	10%	7.09	1, 180	0.008
	59%	67%	8%	4.97	1, 180	0.027
	59%	70%	11%	11.11	1, 180	0.001

Table 7.3. An alternative analysis of the data from trials using different masks (from to: Blank, Faceparts and Noise) represented in the leftmost column.

there was a significant effect of *intensity* ( $F_{3,180} > 62, p < 0.001$ ). For the condition in which the mask was Blank there was a significant interaction between *intensity* and *expression change* ( $F_{3,180} = 3.16, p < 0.05$ ) reflecting heightened accuracy in judging same expression trials when *intensity* = 0%. Since this interaction occurred only in the Blank *mask* condition it may explain the 3-way interaction in the previous analysis. There were no other significant interactions.

The reaction time, the time between a face expression stimulus being displayed and an answer being keyed was recorded. Trials with unusually large reaction time ( $> 1$  standard deviation above average; average = 0.92 seconds,  $sd = 1.05$ ) were removed from the analysis. This removed 7% of trials. A repeated measures ANOVA was performed on reaction time with *expression change* (anti-expression and expression same or different) and *mask type* (Blank, Noise and Faceparts), as within subjects factors, and *intensity* and *delay* as between subjects factors. There was a significant effect of *intensity* ( $F_{3,186} = 6.88, p < 0.001$ ), but no other significant effects or interactions: *expression change* ( $F_{1,186} = 0.76, p = .386$ ), *delay* ( $F_{3,186} = 0.276, p = 0.843$ ) or *mask* ( $F_{2,186} = 2.97, p = 0.054$ ; a trend reflecting longer reaction time with Faceparts, 0.76s, than Blank, 0.73s, or Noise, 0.72s, masks). The

analysis was redone entering all trials and equivalent results were found with only *intensity* having a significant effect.

### **Part 3: Debriefing**

Subjects labelled the anti-expression as corresponding to the expression 20% of the time (which was significantly below chance, 25%, in a one-sample t-test,  $t_{15} = -2.41$ ,  $sd = 0.087$ ,  $p < 0.05$ ) and correctly identified the 50% intensity expressions 87% of the time. No subjects showed any sign of understanding the true nature of the anti-faces during subsequent verbal questioning although many of the subjects did feel features of the faces had been manipulated in some way possibly by mixing together different expressions.

### **Discussion**

Adaptation to an anti-face of a particular expression facilitates subsequent recognition of the corresponding expression in comparison to other expressions. This effect seems to be resistant to changes in the identity of the individual pictured presenting the adapting anti-face expression and the subsequent facial expression for categorisation. The second part of the experiment investigated backwards masking of the anti-expression for delays of up to 1 second before the onset of the facial expression to be identified. It was expected that masking would obliterate the facilitatory effect of adaptation to the anti-face. Instead, it was found that none of the three types of masks obliterated the effect. Longer delays between the disappearance of the anti-face and the onset of the facial expression were expected to significantly diminish the effect. Instead, the insertion of delays of up to a second between the offset of the anti-face and the onset of the facial expression to be categorised had no significant effect. Thus it appears that the adaptation effect is not caused by 'apparent

motion' from the switch between anti-face to the face to be identified. This conclusion concords with the explanation given by Leopold *et al.* (2001).

Leopold *et al.* implicate the STS in their anti-face effect, as neurons in the STS are sensitive to complex groups of features (particularly faces). Some single cell recording studies have shown STS neurons to be robust only to small changes in size (e.g. Ashbridge, Perrett, & Oram, 2000), mirroring the robustness to size changes of 2% found by Leopold *et al.*, while other single cell recording studies (e.g. Rolls & Baylis, 1986) have STS neurons to be robust to large changes in size possibly mirroring another face adaptation effect. Zhao & Chubb (2001) found adaptation after effects for faces that ranged in size by 'two octaves' in a study using face stimuli subjected to horizontal and vertical expansion and contraction (Webster & MacLin, 1999). It might be expected that information from size variant neurons would project to higher level size invariant representations of identity. The results of the previous studies above suggest that this may not be the case; face identity may not be processed in a size invariant manner. This lack of invariance may not be so surprising as identity seems to be processed in a manner that is dependent on different representations for different views (Wallis & Bülthoff, 2001). It would therefore be very interesting to use the methodology used in this chapter to investigate adaptation across facial expression stimuli that differed in size. Leopold *et al.* suggest that in their study adaptation over greater changes in size may not have been possible because large size changes distracted the participants. This may be overcome by the use of masks in between the adaptation and test stimuli.

The enhancement in expression recognition found here caused by adaptation to anti-face images, appears not to be affected by changes in individual identity between adaptation and test stimuli. This may be interpreted as support for the



separate processing of expression from identity posited in Bruce and Young's (1986) model of facial recognition. Although it may be noted that the facilitatory effect of adaptation is numerically smaller when the individual depicted in the anti-expression and expression are different. The smaller adaptation effect for trials in which the individual differs may be due to similar reasons to the diminished identity priming effect found when more dissimilar images of a person are used as prime and target (Johnston & Barry, 2001). It is notable that priming effects bring about faster responses for primed stimuli whereas there was no indication from the results of the current experiment that adaptation to the anti-face might bring about a faster reaction time for recognition corresponding expression.

Desimone (1996) investigated the effects of stimulus repetition on the firing of neurons in the temporal cortex, and comments that the most commonly found response was suppression of firing. Neural responses to stimuli that were similar to the repeated stimulus were suppressed more than responses to less similar stimuli. This may be the mechanism behind the effect shown in the experiment presented here. Adaptation to the anti-face (for a long duration rather than repeatedly) may also cause greatest suppression of response to similar looking stimuli. So, adaptation to anti-fear would suppress responses to anger, sad and disgust stimuli more than to the fear stimuli leading to a greater tendency for fear to be perceived in the subsequent trial. The suppression of firing found by Desimone is relatively long-lived and resistant to the presentation of intervening stimuli in a match to sample task. The suppression is interpreted as a reflection of 'pruning', by suppression, of less functional connections within a neural framework leading to both faster reaction times (e.g. with priming) and lower levels of neural activity when a similar stimulus is presented (Grill-Spector, et al., 1999). This interpretation is in line with the adaptation effect found here being

resilient to masking and the delays of up to a second tested. Other studies have found that masking interrupts neuronal firing in the STS but were not designed to investigate whether masking would disrupt suppression (Keysers & Perrett, 2002; Kovács, Vogels, & Orban, 1995; Rolls, Toveé, & Panzeri, 1999).

The effect of adaptation to anti-expressions may, alternatively, be due to suppression of higher-level connections that are more specific to emotion processing or be mediated through a non-striate cortex route. If so, then it might be expected that the effect size would differentiate among emotions. The experiment presented here was not designed to investigate this, although it is notable that the adaptation effect tends to occur more for fear than for other emotions (difference in proportion of expressions recognised between corresponding and non-corresponding conditions in Part 2: disgust 7%, anger 8%, sad 12%, and fear 15%) possibly suggesting a role for the amygdala, a structure that has been noted to be responsive facial expressions, particularly of fear, and to use the colliculo-pulvinar pathway, a possible alternative route for visual recognition, when stimuli are repeatedly presented for very short time intervals using masks (Morris, Öhman, & Dolan, 1999; Whalen, et al., 1998) or when striate cortex is not present (de Gelder, Vroomen, Pourtois, & Weiskrantz, 1999).

In summary, adaptation to anti-expressions causes subsequently presented expressions to be recognised more readily by subjects as the corresponding expression. This adaptation effect is robust to masking with a variety of masks and is not significantly affected by delays between the offset of the anti-expression and the onset of the expression of up to 1 second. These qualities and other evidence presented previously (Leopold, O'Toole, Vetter, & Blanz, 2001) suggest that neurons in the STS may be suppressed by adapting to the anti-expressions. In the experiment reported here, suppression would be expected to affect representations corresponding

to the facial expression that is most dissimilar to the anti-face least, and the more similar expression the most. So, adaptation to an anti-fear face would suppress representations related to fear the least, as the fear face is less similar to the anti-expression than any of the other facial expression.

## **Chapter 8: Thesis discussion.**

Over the past few years advances in computer technology have enabled more realistic graphic manipulation for face perception research. Developments in computer graphics have made the study of different areas of face perception possible and have lead to more realistic face stimuli of greater ecological validity. In this thesis a number of different techniques for the manipulation of face images were reviewed, and selected techniques applied to study different areas of face perception. Here an overview is presented in which results are summarized, and several links between the studies and avenues for further development are highlighted.

In the second chapter a novel method was used to produce chimaeric face stimuli for the investigation of an effect first studied by Wolff (1933), namely, that the left side of faces (from the viewer's perspective) seems more representative of an individual than the right. The use of a new computer graphics technique meant that stimuli could be made in which there was no discernable midline between the pairs of face halves used in stimulus manufacture. This is important, as the presence of a join in chimaeric face stimuli has been shown to affect perception (Milner & Dunne, 1977). Apart from the use of more realistic stimuli, this experiment also advanced the literature by demonstrating that under free fixation, left of stimulus attentional biases were also present for attractiveness judgements and right side of stimulus attentional biases were present for lip reading. The significance level of this right of stimulus attentional bias was not very high ( $p < 0.05$ ; 58% choosing the right side of the stimulus), although the same effect was found for both of the lip-reading stimuli. These effects may be caused by asymmetrical differences in eye-scanning patterns and eye tracking equipment would thus be useful to explore the bias. It is notable that a recent electrophysiological study (McCarthy, Puce, Belger, & Allison, 1999) has also suggested a right hemi-field bias for interpreting the lips. The use of still images

in this thesis may not employ normal lip-reading processes fully. Therefore in further work it would be advantageous to examine perceptual bias in the viewing of animated (speaking) chimaeric faces. The investigation of the lateral biases in perception of the McGurk effect, in which the phoneme heard is altered by the mouth movement seen, would be especially interesting as this effect is related to the combination of visual and auditory information.

Chimaeric stimuli that differ between their left and right hand sides could also be used to study the adaptation effect found in Chapter 7. Adaptation effects might be expected to vary depending on the differences between the two hemispheres. Differences between the hemispheres may be found at several levels; for example, the right hemisphere has been implicated in configural coding whereas the left hemisphere plays a role in feature coding. Thus, alterations in the configuration of the anti-expression stimuli might be expected to degrade the facilitating effect of the anti-expressions more for the right than the left hemisphere, whereas alterations to particular features might be hypothesised to degrade the effect related to the left hemisphere. Differences might also be expected depending upon the facial trait being assessed. For example it might be expected that the strength of the adaptation effect would mirror the laterality biases seen in Chapter 2, with all face dimensions, except lip-reading, being lateralised to the left of stimulus, implicating the right hemisphere.

Chapter 3 presented novel methods to make symmetrical versions of faces. Previous research had suggested that in many animals, including humans, symmetry is an indicator of health and that symmetrical individuals are preferred as mates since healthy looking individuals also tend to be more symmetrical. The preference for more symmetrical individuals may be due to selection of other cues to health rather than symmetry itself. Studies were therefore performed in which the effect on



attractiveness was examined when symmetry was directly manipulated, for example by removal of plumage in either a symmetrical or asymmetrical patterns (Swaddle & Cuthill, 1995). Non-human animals were found, in the majority of studies, to prefer the more symmetrical mate showing that symmetry is indeed preferred. In humans, by contrast, manipulations of symmetry performed using graphic methods had led to the conclusion that we prefer less symmetrical individuals (Kowner, 1996, 1997; Langlois, Roggman and Musselman, 1994; Swaddle & Cuthill, 1995). Experiments presented in Chapter 3, using more appropriate computer graphic methods, showed that this conclusion was not correct and that humans do find symmetrical faces more attractive. Chapter 3 then presented two investigations into factors that may modulate strength of symmetry preference. The first of these studies found that women who rate themselves as being more attractive are more selective concerning facial symmetry, choosing a greater proportion of symmetrical male faces when judging facial attractiveness. This greater selectivity may relate to being better able to gain and keep an attractive mate or being better able to tolerate losing a mate. In order to investigate this further another experiment was performed in which BDI score was related to selectivity for symmetry. In student populations that have not been diagnosed as having depression, higher BDI levels have been found to be highly correlated to self-report scores of anxiety and have been suggested to reflect a general malaise (Gotlib, 1984). It was expected that low BDI score, taken to indicate lower levels of this malaise, would therefore relate to greater selectivity for symmetry. Instead, the opposite was true. Higher BDI score related to higher selectivity. This may possibly reflect disappointment with relationships, although, further research is needed to understand this relationship. Such research could use stimuli that trade off different cues to genetic fitness, like symmetry, against other attractive cues, like smiling. The

use of this sort of trade-off would enable strength of preference for fitness cues to be assessed rather than sensitivity to fitness cues.

In Chapter 4, colour caricaturing was used in a novel manner to aid in the visualization of differences related to skin health. Transformation of skin colouration of the images of different individuals was then shown manipulate colour cues that relate to perceived skin health. These captured colour cues may in the future be used to examine individual differences between people in their perception of skin health and may be used to examine the different health factors that are valued for different types of relationship.

In Chapter 5, a foray into dynamic cues influencing face perception was presented. The experiment demonstrated that aspects of tasting behaviour provide dynamic cues to how the different tastes can be perceived. The experiment lays down some of the psychological groundwork for the future capture and manipulation of dynamic facial actions related to taste perception. Although, it is now possible to transform dynamically moving faces (Tiddeman & Perrett, 2001), currently the techniques have not been developed to prototype and manipulate dynamic actions made by faces when the time course of the actions varies between exemplars.

In Chapter 6, a novel computer graphics method was used to make anti-expressions. Anti-expressions are the opposite in colouration, texture and configuration of facial expressions with regards to the neutral expression. So, for example, in a happy expression the ends of the lips are turned up in comparison to neutral, in the anti-happy expression they are turned down. Using these stimuli, a dynamic test was produced with the aim of assessing, in terms of the 6 basic emotional expressions, what people perceived as a neutral facial expression. The test was run on depressed and non-depressed subjects and it was found that the depressed

subjects' percept of what is a neutral facial expression was different from that of non-depressed individuals. When asked to select a "neutral" expression, depressed subjects tended to choose higher intensities of disgust and anger as being neutral than non-depressed subjects. In other words, depressed subjects' percept of neutral tends to be more disgusted and angry than non-depressed subjects' percept. This work on the perception of neutral facial expressions lays a basis for understanding how individuals differ in the way they perceive a neutral face.

The last experimental chapter, Chapter 7, investigated adaptation to the anti-face expression stimuli used in the pervious chapter. Adaptation to anti-expressions was found to result in a subsequent facilitation of recognition of the corresponding expression. For example, adaptation to anti-fear results in a subsequently displayed fear face as being more likely to be labelled as looking "fearful". The effect was found to generalize across face stimulus identity and to be robust to both masking and to delays of up to a second between the presentation of the adapting stimulus and the stimulus to be identified by the subject. The effects found in this experiment were interpreted as being a result of the suppression of neurons selective for facial configuration in the temporal cortex. Further work may determine the location of suppressed neurons and shed light on what representation is being adapted. For example, adaptations may be at (or above) the level of groups of features coding for the facial expression across the whole face. Adaptation to one side of the face might then be expected to affect the perception of the other side of the face. So, if the right side anti-face expression is adapted to and then the left side of the face is subsequently presented for identification then, an effect would still be present. Alternatively, it is possible that adaptation occurs at a semantic level in which case an analogous effect

may be found for emotional words when the corresponding anti-emotional face is adapted.

## ***Summary***

This thesis presented a number of computer graphic techniques for use in the manipulation of face image stimuli in face perception research. These techniques were shown to be applicable to the investigation of possible hemispheric asymmetries in perception of facial qualities, the role of facial symmetry in attractiveness judgements, perceived skin health, dynamic cues related to food consumption and the perception of facial expressions. These different areas of facial research are ongoing, and although presented as separate chapters examining different ideas they have all benefited from the cross fertilization of ideas and techniques.

## References

- Adolphs, R., Tranel, D., Damasio, H., & Damasio, A. (1994). Impaired recognition of emotion in facial expressions following bilateral damage to the human amygdala. *Nature*, 372, 699-672.
- Adolphs, R., Tranel, D., Damasio, H., & Damasio, A. (1995). Fear and the human amygdala. *Journal of Neuroscience*, 15, 5879-5891.
- Ashbridge, E., Perrett, D. I., & Oram, M. W. (2000). Effect of image orientation and size on object recognition: Responses of single units in the macaque monkey temporal cortex. *Cognitive Neuropsychology*, 17, 13-34.
- Bailit, H., Workman, P., Niswander, J., & Maclean, J. (1970). Dental asymmetry as an indicator of genetic and environmental conditions in human populations. *Human Biology*, 42, 626-638.
- Baker, E., & Seltzer, M. (1994). *Evolving line drawings*. Paper presented at the Fifth Annual Conference on Genetic Algorithms, Banff, Alberta, Canada.
- Bakker, T., Kunzler, R., & Mazzi, D. (1999). Sexual selection - Condition-related mate choice in sticklebacks. *Nature*, 401, 234.
- Bateman, A. J. (1983). *Mate choice*. London: Cambridge University Press.
- Beale, J. L., & Keil, F. C. (1995). Categorical effects in the perception of faces. *Cognition*, 57, 217-239.
- Beck, A. T., Rush, A., Shaw, B., & Emery, G. (1979). *Cognitive Therapy of Depression*. New York: John Wiley and Sons / Guilford Press.
- Beck, A. T., & Steer, R. A. (1987). *Beck Depression Inventory Manual*. San Antonio: The Psychological Corporation Harcourt Brace Jovanovich, Inc.
- Beck, A. T., Ward, C., Mendelson, M., Mock, J., & Erbaugh, J. (1961). An inventory for measuring depression. *Archives of General Psychiatry*, 4, 53-63.
- Beier, T., & Neely, S. (1992). Feature-based image metamorphosis. *Computer Graphics*, 26, 35-42.
- Benson, P. J. (1992). *Perception and recognition of computer-enhanced facial attributes and abstracted prototypes*. Unpublished Phd Thesis, University of St Andrews, St Andrews, UK.

- Benson, P. J., & Perrett, D. I. (1991a). Computer averaging and manipulation of faces. In P. Wombell (Ed.), *Photovideo: Photography in the Age of the Computer* (pp. 32-38). London: Rivers Oram Press.
- Benson, P. J., & Perrett, D. I. (1991b). Perception and recognition of photographic quality caricatures: implications for the recognition of natural images. *European Journal of Cognitive Psychology*, 3, 105-135.
- Benson, P. J., & Perrett, D. I. (1991c). Synthesising continuous-tone caricatures. *Image and Vision Computing*, 9, 123-129.
- Benson, P. J., & Perrett, D. I. (1993). Extracting prototypical facial images from exemplars. *Perception*, 22, 257-262.
- Benson, P. J., & Perrett, D. I. (1994). Visual processing of facial distinctiveness. *Perception*, 23, 75-93.
- Berenbaum, H. (1992). Posed facial expressions of emotion in schizophrenia and depression. *Psychological Medicine*, 22, 929-937.
- Berridge, K. (2000). Measuring hedonic impact in animals and infants: microstructure of affective taste reactivity patterns. *Neuroscience and Biobehavioural Reviews*, 24, 173-189.
- Berry, D. S. (1990). What can a moving face tell us? *Journal of Personality and Social Psychology*, 58, 1004-1014.
- Berry, S. B., & McArthur, L. Z. (1986). Perceiving character in faces: The impact of age-related craniofacial changes on social perception. *Psychological Bulletin*, 100, 3-18.
- Beymer, D., Shashua, A., & Poggio, T. (1993). Example based image analysis and synthesis. *M.I.T. Artificial Intelligence Laboratory, A.I. Memo*, no. 1431.
- Blair, R. J. R., & Curran, H. V. (1999). Selective impairment in the recognition of anger induced by diazepam. *Psychopharmacology*, 147, 335-338.
- Blair, R. J. R., Morris, J. S., Frith, C. D., Perrett, D. I., & Dolan, R. J. (1999). Dissociable neural responses to facial expressions of sadness and anger. *Brain*, 122, 883-893.
- Blanz, V., O'Toole, A. J., Vetter, T., & Wild, H. A. (2000). On the other side of the mean: the perception of dissimilarity in human faces. *Perception*, 29, 885-891.
- Blanz, V., & Vetter, T. (1999). *A morphable model for the synthesis of 3D faces*. Los Angeles: Addison Wesley Longman.



- Booth, A., & Dabbs, J. (1993). Testosterone and men's marriages. *Social Forces*, 72, 463-477.
- Bornstein, R. F. (1989). Exposure and affect: overview and meta-analysis of research 1968-1987. *Psychological Bulletin*, 106, 265-289.
- Borod, J. C., Caron, H. S., & Koff, E. (1981). Asymmetry of facial expression related to handedness, footedness, and eyedness: a quantitative study. *Cortex*, 17, 381-390.
- Borod, J. C., & Koff, E. (1983). Hemiface mobility and facial expression asymmetry. *Cortex*, 19, 327-332.
- Bouhuys, A. L., Bloem, G. M., & Groothuis, T. G. G. (1995). Induction of depressed and elated mood by music influences the perception of facial emotional expressions in healthy subjects. *Journal of Affective Disorders*, 33, 215-226.
- Bouhuys, A. L., Geerts, E., & Gordijn, M. C. M. (1999a). Depressed patients' perceptions of facial emotions in depressed and remitted states are associated with relapse - A longitudinal study. *Journal of Nervous and Mental Disease*, 187, 595-602.
- Bouhuys, A. L., Geerts, E., & Gordijn, M. C. M. (1999b). Gender-specific mechanisms associated with outcome of depression: perception of emotions, coping and interpersonal functioning. *Psychiatry Research*, 85, 247-261.
- Bouhuys, A. L., Geerts, E., Mersch, P. P. A., & Jenner, J. A. (1996). Nonverbal interpersonal sensitivity and persistence of depression: perception of emotions in schematic faces. *Psychiatry Research*, 64, 193-203.
- Bouhuys, A. L., Geerts, E., & Mersch, P. (1997). Relationship between perception of facial emotions and anxiety in clinical depression: does anxiety-related perception predict persistence of depression? *Journal of Affective Disorders*, 43, 213-223.
- Brennan, S. E. (1982). *Caricature generator*. Unpublished masters thesis, MIT.
- Brennan, S. E. (1985). The caricature generator. *Leonardo*, 18, 170-187.
- Bruce, V., & Young, A. (1986). Understanding face recognition. *British Journal of Psychology*, 77, 305-327.
- Bruyer, R., & Granato, P. (1999). Categorical effects in the perception of facial expressions: MARIE a simple and discriminating clinical tool. *European Review of Applied Psychology*, 49, 3-10.

- Burke, P. H. (1971). Stereophotogrammetric measurement of normal facial asymmetry in children. *Human Biology*, 43, 536-548.
- Burt, D. M., & Perrett, D. I. (1995). Perception of age in adult Caucasian male faces - computer graphic manipulation of shape and color information. *Proceedings of the Royal Society of London Series B-Biological Sciences*, 259, 137-143.
- Burt, D. M., & Perrett, D. I. (1997). Perceptual asymmetries in judgements of facial attractiveness, age, gender, speech and expression. *Neuropsychologia*, 35, 685-693.
- Calder, A. J., Rowland, D., Young, A. W., Nimmo-Smith, I., Keane, J., & Perrett, D. I. (2000). Caricaturing facial expressions. *Cognition*, 76, 105-146.
- Calder, A. J., Young, A. W., Perrett, D. I., Etcoff, N. L., & Rowland, D. (1996a). Categorical perception of morphed facial expressions. *Visual Cognition*, 3, 81-117.
- Calder, A. J., Young, A. W., Rowland, D., & Perrett, D. I. (1997). Computer-enhanced emotion in facial expressions. *Proceedings of the Royal Society of London Series B-Biological Sciences*, 264, 919-925.
- Calder, A. J., Young, A. W., Rowland, D., Perrett, D. I., Hodges, J. R., & Etcoff, N. L. (1996b). Facial emotion recognition after bilateral amygdala damage - differentially severe impairment of fear. *Cognitive Neuropsychology*, 13, 699-745.
- Caldwell, C., & Johnston, V. S. (1991). *Tracking a criminal suspect through "face-space" with a genetic algorithm*. Paper presented at the Forth International Conference on Genetic Algorithms, San Mateo.
- Campanella, S., Hanoteau, C., Depy, D., Rossion, B., Bruyer, R., Crommelinck, M., & Guerit, J. (2000). Right N170 modulation in a face discrimination task: An account for categorical perception of familiar faces. *Psychophysiology*, 37, 796-806.
- Campbell, R. (1986). The lateralization of lip-read sounds: a first look. *Brain and Cognition*, 5, 1-21.
- Campbell, R. (1987). The cerebral lateralization of lip-reading. In B. Dodd & R. Campbell (Eds.), *Hearing by Eye: the Psychology of Lip-Reading* (pp. 215-226). London: Laurence Album Associates.
- Christman, S. D., & Hackworth, M. D. (1993). Equivalent perceptual asymmetries for free viewing of positive and negative emotional expressions in chimaeric faces. *Neuropsychologia*, 31, 621-624.

- Cooley, E. L., & Nowicki, S. (1989). Discrimination of facial expressions of emotion by depressed subjects. *Genetic, Social and General Psychology Monographs*, 115, 451-465.
- Cox, B. J., Enns, M. W., & Larsen, D. K. (2001). The continuity of depression symptoms: Use of cluster analysis for profile identification in patient and student samples. *Journal of Affective Disorders*, 65, 67-73.
- Coyne, J. C. (1976). Depression and the response of others. *Journal of Abnormal Psychology*, 85, 186-193.
- Cunningham, M. R. (1986). Measuring the physical in physical attractiveness - quasi-experiments on the sociobiology of female facial beauty. *Journal of Personality and Social Psychology*, 50, 925-935.
- Cunningham, M. R., Barbee, A. P., & Pike, C. L. (1990). What do women want? Facialmetric assessment of multiple motives in the perception of male physical attractiveness. *Journal of Personality and Social Psychology*, 59, 61-72.
- Daugman, J. G. (1980). Two-dimensional spectral analysis of cortical receptive field profiles. *Vision Research*, 20, 847-56.
- de Gelder, B., Vroomen, J., Pourtois, G., & Weiskrantz, L. (1999). Non-conscious recognition of affect in the absence of striate cortex. *NeuroReport*, 10, 3759-3763.
- De Renzi, E., Garthia, M. B., & Faglioni, P. (1989). Right posterior brain damaged patients are poor at assessing the age of a face. *Neuropsychologia*, 27, 839-848.
- Deffenbacher, K., Vetter, T., Johanson, J., & O'Tool, A. (1998). Facial aging, attractiveness and distinctiveness. *Perception*, 27, 1233-1243.
- Desimone, R. (1996). Neural mechanisms for visual memory and their role in attention. *Proceeding of the National Academy of Sciences*, 93, 13494-13499.
- Dion, K., Berscheid, E., & Walster, E. (1972). What is beautiful is good. *Journal of Personality and Social Psychology*, 24, 285-290.
- Ekman, P., & Friesen, W. V. (1976). *Pictures of Facial Affect*. Palo Alto, California: Consulting Psychological Press.
- Enquist, M., & Arak, A. (1994). Symmetry, beauty and evolution. *Nature*, 372, 169-172.
- Enquist, M., & Johnstone, R. A. (1997). Generalization and the evolution of symmetry preferences. *Proceedings of the Royal Society of London Series B-Biological Sciences*, 264, 1345-1348.

- Etcoff, N. L., & Magee, J. J. (1992). Categorical perception of facial expressions. *Cognition*, 44, 227-240.
- Feingold, A. (1988). Matching for attractiveness in romantic partners and same-sex friends - a meta-analysis and theoretical critique. *Psychological Bulletin*, 104, 226-235.
- Feingold, A. (1992). Good-looking people are not what we think. *Psychological Bulletin*, 111, 304 - 341.
- Ferrario, V. F., Sforza, C., Ciusa, V., Dellavia, C., & Tartaglia, G. M. (2001). The effect of sex and age on facial asymmetry in healthy subjects: A cross-sectional study from adolescence to mid-adulthood. *Journal of Oral and Maxillofacial Surgery*, 59, 382-388.
- Fisher, R. A. (1930). *The genetical theory of natural selection*. Oxford: Clarendon Press.
- Frigerio, E., Burt, D. M., Peverelli, M., Borgatti, R., Gagliardi, C., Montirosso, R., Pozzoli, S., Cazzaniga, I., & Perrett, D. I. (2001). *Perception of facial affect in children with cognitive difficulties caused by Williams and Down's syndromes*. Paper presented at the FACS 2001 - 9th European Conference Facial Expression, Measurement and Meaning, Innsbruck, Austria.
- Frowd, C. D. (2000). *EvoFIT: A Holistic, Evolutionary Facial Imaging System*, [WWW]. Available: <http://www.stir.ac.uk/Departments/HumanSciences/Psychology/Postgrads/cdf1/EvoFIT/> [2001, January].
- Fryday, S. L., & Greigsmith, P. W. (1994). The effects of social-learning on the food choice of the house sparrow (*Passer-domesticus*). *Behaviour*, 128, 281-300.
- Furlow, F. B., Armijo-Prewitt, T., Gangestad, S. W., & Thornhill, R. (1997). Fluctuating asymmetry and psychometric intelligence. *Proceedings of the Royal Society of London Series B-Biological Sciences*, 264, 823-829.
- Gagliardi, C., Frigerio, E., Cazzaniga, I., Burt, D. M., Peverelli, M., Borgatti, R., & Perrett, D. I. (2001). *Using computer for studying emotion recognition in children with cognitive impairment: methodological aspects*. Paper presented at the Health Psychology 2001 conference, St Andrews, UK.
- Galef, B. G. (1995). Food selection: problems in understanding how we choose foods to eat. *Neuroscience and Biobehavioural Reviews*, 20, 67-73.
- Gallese, V., Fadiga, L., Fogassi, L., & Rizzolatti, G. (1996). Action recognition in the premotor cortex. *Brain*, 119, 593-609.

- Galton, F. J. (1878). Composite portraits. *Nature*, 18, 97-100.
- Gangestad, S. W., Bennett, K. L., & Thornhill, R. (2001). A latent variable model of developmental instability in relation to men's sexual behaviour. *Proceedings of the Royal Society of London Series B-Biological Sciences*, 268, 1677-1684.
- Gangestad, S. W., & Buss, D. M. (1993). Pathogen prevalence and human mate preferences. *Ethology and Sociobiology*, 14, 89-96.
- Gangestad, S. W., & Thornhill, R. (1998). Menstrual cycle variation in women's preferences for the scent of symmetrical men. *Proceedings of the Royal Society of London Series B-Biological Sciences*, 265, 927-933.
- Gangestad, S. W., Thornhill, R., & Yeo, R. A. (1994). Facial attractiveness, developmental stability, and fluctuating asymmetry. *Ethology and Sociobiology*, 15, 73-85.
- George, M. S., Huggins, T., McDermut, W., Parekh, P. I., Rubinow, D., & Post, R. M. (1998). Abnormal facial emotion recognition in depression: serial testing in an ultra-rapid-cycling patient. *Behavior Modification*, 22, 192-204.
- Gessler, S., Cutting, J., Frith, C. D., & Weinman, J. (1989). Schizophrenic inability to judge facial emotion: a controlled study. *British Journal of Clinical Psychology*, 28, 19-29.
- Gilbert, A. N., Fridlund, A. J., & Sabini, J. (1987). Hedonic and social determinants of facial displays to odors. *Chemical Senses*, 12, 355-363.
- Gilbert, C., & Bakan, P. (1973). Visual asymmetry in perception of faces. *Neuropsychologia*, 11, 355-362.
- Gomi, A. (1988). *Americans 1.0 Los Angeles 1994* [CD-ROM]. Bunkasha, Japan: Digitalogue.
- Gotlib, I. H. (1984). Depression and general psychopathology in university students. *Journal of Abnormal Psychology*, 93, 19-30.
- Gotlib, I. H., & Robinson, L. A. (1982). Responses to depressed individuals: discrepancies between self-report and observer-rated behavior. *Journal of Abnormal Psychology*, 91, 231-240.
- Grammer, K., & Thornhill, R. (1994). Human (Homo-Sapiens) facial attractiveness and sexual selection - the role of symmetry and averageness. *Journal of Comparative Psychology*, 108, 233-242.
- Graves, R., Goodglass, H., & Landis, T. (1982). Mouth asymmetry during spontaneous speech. *Neuropsychologia*, 20, 371-381.

- Graves, R. E., Strauss, E. H., & Wada, J. (1990). Mouth asymmetry during speech of epileptic patients who have undergone carotid amytal testing. *Neuropsychologia*, 28, 1117-1121.
- Griffiths, R. W., & Kunz, P. R. (1973). Assortative mating: a study of physiognomic homogamy. *Social Biology*, 20, 448-453.
- Grill-Spector, K., Kushnir, T., Edelman, S., Avidan, G., Itzhak, Y., & Malach, R. (1999). Differential processing of objects under various viewing conditions in the human lateral occipital complex. *Neuron*, 24, 187-203.
- Gur, R. C., Erwin, R. J., Gur, R. E., Zwi, A. S., Heimber, C., & Kraemer, H. C. (1992). Facial emotion discrimination: II. Behavioral findings in depression. *Psychiatry Research*, 42, 241-251.
- Hagen, M., & Perkins, D. (1983). A refutation of the hypothesis of the superfidelity of caricatures relative to photographs. *Perception*, 12, 55-61.
- Hale, W. W. (1998). Judgment of facial expressions and depression persistence. *Psychiatry Research*, 80, 265-274.
- Hancock, P. J. B. (2000). Evolving faces from principal components. *Behaviour Research Methods, Instruments and Computers*, 32, 327-333.
- Hancock, P. J. B., Bruce, V., & Burton, A. M. (1998). A comparison of two computer-based face recognition systems with human perceptions of faces. *Vision Research*, 38, 2277-2288.
- Hancock, P. J. B., Burton, A. M., & Bruce, V. (1996). Face processing: human perception and principal components analysis. *Memory and Cognition*, 24, 26-40.
- Hancock, P. J. B., & Frowd, C. D. (1999, 8th and 9th of April 1999). *Evolutionary generation of faces*. Paper presented at the Symposium on AI and Scientific Creativity, Edinburgh.
- Harnad, S. (1987). *Categorical Perception*. Cambridge, UK: Cambridge University Press.
- Hasselmo, M. E., Rolls, D. I., & Baylis, G. C. (1989). The role of expression and identity in the face-selective responses of neurons in the temporal visual cortex of the monkey. *Behavioural Brain Research*, 32, 203-218.
- Hawkins, S. S., Perrett, D. I., Burt, D. M., Rowland, D. A., & Murahata, R. I. (1999). Prototypes of facial attributes developed through image averaging techniques. *International Journal of Cosmetic Science*, 21, 159-166.



- Henss, R. (1991). Perceiving age and attractiveness and facial photographs. *Journal of Applied Social Psychology*, 21, 933-946.
- Hickok, G., & Poeppel, D. (2000). Towards a functional neuroanatomy of speech perception. *Trends in Cognitive Sciences*, 4, 131-138.
- Hishimura, Y. (2000). Enhancement of food aversion by exposure to a poisoned conspecific in Norway rats (*Rattus norvegicus*). *Japanese Psychological Research*, 42, 183-187.
- Holland, J. H. (1992). Genetic Algorithms. *Scientific American*, July, 44-50.
- Humphreys, G. W., Donnelly, N., & Riddoch, M. J. (1993). Expression is computed separately from facial identity, and it is computed separately for moving and static faces - neurophysiological evidence. *Neuropsychologia*, 31, 173-181.
- Hurlbert, A. (2001). Trading faces. *Nature Neuroscience*, 4, 3-5.
- Hwang, B., Blanz, V., Vetter, T., & Lee, S. (2000). *Face reconstruction from a small number of feature points*. Paper presented at the Proceedings of the First IEEE International Workshop, BMCV 2000.
- Isono, K., Oda, M., & Akamatsu, S. (1997). Generating subjective impression for the face using genetic algorithms. *Technical Report of the IEICE, HIP*, 97-4.
- Jancke, L., & Kaufmann, N. (1994). Facial EMG responses to odors in solitude and with an audience. *Chemical Senses*, 19, 99-111.
- Johnston, R. A., & Barry, C. (2001). Best face forward: similarity effects in repetition priming of face recognition. *The Quarterly Journal of Experimental Psychology*, 54A, 383-396.
- Johnston, V. S. (1994). Methods and apparatus for generating composites of human faces. *US Patent*, 5,375,195, December 20, 1994.
- Johnston, V. S., & Franklin, M. (1993). Is beauty in the eye of the beholder? *Ethology and Sociobiology*, 14, 183-199.
- Johnston, V. S., Hagel, R., Franklin, M., Fink, B., & Grammer, K. (2001). Male facial attractiveness - evidence for hormone-mediated adaptive design. *Evolution and Human Behavior*, 22, 251-267.
- Jones, B. C., Little, A. C., Penton-Voak, I. S., Tiddeman, B. P., Burt, D. M., & Perrett, D. I. (2001). Facial symmetry and judgements of apparent health: Support for a "good genes" explanation of the attractiveness-symmetry relationship. *Evolution and Human Behavior*, 22, 417-429.

- Kamachi, M., Bruce, V., Mukaida, S., Gyoba, J., Yoshikawa, S., & Akamatsu, S. (2001). Dynamic properties influence the perception of facial expressions. *Perception*, 30, 875-887.
- Keating, C. F. (2001). Charismatic faces: social status cues put face appeal in context. In G. Rhodes & L. Zebrowitz (Eds.), *Facial Attractiveness: Evolutionary, Cognitive and Social Perspectives* (Vol. 1, ). Westport, Connecticut: Ablex.
- Keles, P., Diyarbakirli, S., Tan, M., & Tan, U. (1997). Facial asymmetry in right- and left-handed men and women. *International Journal of Neuroscience*, 91, 147-159.
- Keysers, C., & Perrett, D. I. (2002). Visual masking and RSVP reveal neural competition. *Trends in Cognitive Sciences*, 6, 120-125.
- Kircher, T. T. J., Senior, C., Phillips, M. L., Benson, P. J., Bullmore, E. T., Brammer, M., Simmons, A., Williams, S., Bartels, M., & David, A. S. (2000). Towards a functional neuroanatomy of self processing: effects of faces and words. *Cognitive Brain Research*, 10, 133-144.
- Kircher, T. T. J., Senior, C., Phillips, M. L., Rabe-Hesketh, S., Benson, P. J., Bullmore, E. T., Brammer, M., Simmons, A., Bartels, M., & David, A. S. (2001). Recognizing one's own face. *Cognition*, 78, 1-15.
- Kotsoni, E., de Haan, M., & Johnson, M. H. (2001). Categorical perception of facial expressions by 7-month-old infants. *Perception*, 30, 1115-1125.
- Kovács, G., Vogels, R., & Orban, G. A. (1995). Cortical correlates of pattern backward masking. *Proceedings of the National Academy of Science*, 92, 5587-5591.
- Kowner, R. (1995). Laterality and facial expression and its effect on attributions of emotion and personality: a reconsideration. *Neuropsychologia*, 33, 539-559.
- Kowner, R. (1996). Facial asymmetry and attractiveness judgement in developmental perspective. *Journal of Experimental Psychology: Human Perception and Performance*, 22, 662-675.
- Kowner, R. (1997). The perception and attribution of facial asymmetry in normal adults. *Psychological Record*, 47, 371-384.
- Kowner, R. (2001). Psychological perspective on human developmental stability and fluctuating asymmetry: Sources, applications and implications. *British Journal of Psychology*, 92, 447-469.
- Langlois, J. H., Kalakanis, L., Rubenstein, A. J., Larson, A., Hallam, M., & Smoot, M. (2000). Maxims or myths of beauty? A meta-analytic and theoretical review. *Psychological Bulletin*, 126, 390 - 423.

- Langlois, J. H., & Roggman, L. A. (1990). Attractive faces are only average. *Psychological Science*, 1, 115-121.
- Langlois, J. H., Roggman, L. A., & Musselman, L. (1994). What is average and what is not average about attractive faces. *Psychological Science*, 5, 214-220.
- Langlois, J. H., Roggman, L. A., & Rieser-Danner, L. A. (1990). Infants' differential social responses to attractive and unattractive faces. *Developmental Psychology*, 26, 153-159.
- Lee, K. J. (1997). *Digital manipulation of faces and its consequences upon identification and attractiveness*. Unpublished Phd Thesis, University of St Andrews, St Andrews, UK.
- Lee, K. J., Byatt, G., & Rhodes, G. (2000). Caricature effects, distinctiveness, and identification: testing the face-space framework. *Psychological Science*, 11, 379-385.
- Lee, K. J., & Perrett, D. I. (1997). Presentation-time measures of the effects of manipulations in colour space on discrimination of famous faces. *Perception*, 26, 733-752.
- Lens, L., & Van Dongen, S. (2000). Fluctuating and directional asymmetry in natural bird populations exposed to different levels of habitat disturbance, as revealed by mixture analysis. *Ecology Letters*, 3, 516-522.
- Leopold, D. A., O'Toole, A. J., Vetter, T., & Blanz, V. (2001). Prototype-referenced shape encoding revealed by high-level aftereffects. *Nature Neuroscience*, 4, 89-94.
- Leung, B., & Forbes, M. R. (1996). Fluctuating asymmetry in relation to stress and fitness: effects of trait type as revealed by meta-analysis. *Ecoscience*, 3, 400-413.
- Lewinsohn, P. M. (1975). The behavioral study and treatment of depression. In M. Hersen & R. Eisler (Eds.), *Progress in Behavior Modification* (pp. 19-64). NY: Academic Press.
- Little, A. C., Burt, D. M., Penton-Voak, I. S., & Perrett, D. I. (2001). Self-perceived attractiveness influences human female preferences for sexual dimorphism and symmetry in male faces. *Proceedings of the Royal Society of London Series B-Biological Sciences*, 268, 39-44.
- Little, A. C., & Hancock, P. J. (In press). The role of masculinity and distinctiveness on the perception of attractiveness in human male faces. *British Journal of Psychology*.

- Lopez, S. (1999). Parasitized female guppies do not prefer showy males. *Animal Behaviour*, 57, 1129-1134.
- Luh, K. E., Redl, J., & Levy, J. (1994). Left- and right-handers see people differently: free-vision perceptual asymmetries for chimaeric stimuli. *Brain and Cognition*, 25, 141-160.
- Malina, R. M., & Buschang, P. H. (1984). Anthropometrics symmetry in normal and mentally retarded males. *Annals of Human Biology*, 11, 515-531.
- Manning, J. T., & Chamberlain, A. T. (1993). Fluctuating asymmetry, sexual selection and canine teeth in primates. *Proceedings of the Royal Society of London Series B-Biological Sciences*, 251, 83-87.
- Marks, P. J., Vipond, I., Carlisle, D., Deakin, D., Fey, R. E., & Caul, E. O. (2000). Evidence for airborne transmission of Norwalk-like virus (NLV) in a hotel restaurant. *Epidemiology and Infection*, 124, 481-487.
- Mason, J. R., Arzt, A. H., & Reidinger, R. F. (1984). Comparative assessment of food preferences and aversions acquired by blackbirds via observational learning. *Auk*, 101, 796-803.
- Mason, J. R., & Reidinger, R. F. (1982). Observational learning of food aversions in Red-winged Blackbirds (*Agelaius phoeniceus*). *Auk*, 99, 548-554.
- McCarthy, G., Puce, A., Belger, A., & Allison, T. (1999). Electrophysiological studies of human face perception. II: response properties of face-specific potentials generated in occipitotemporal cortex. *Cerebral cortex*, 5, 431-444.
- Mealey, L., Bridgestock, R., & Townsend, G. (1999). Symmetry and perceived facial attractiveness. *Journal of Personality and Social Psychology*, 76, 151-158.
- Micrografx. (1987-1999). IGrafx Image 1/Picture Publisher: Micrografx Inc.
- Millinski, M., & Bakker, T. C. (1990). Female sticklebacks use male coloration in sticklebacks and therefore avoid parasitised males. *Nature*, 344, 330-333.
- Milner, A. D., & Dunne, J. J. (1977). Lateralised perception of bilateral chimaeric faces by normal subjects. *Nature*, 268, 175-176.
- Møller, A. P. (1990). Fluctuating asymmetry in male sexual ornaments may reliably reveal male quality. *Animal Behavior*, 40, 1185-1187.
- Møller, A. P. (1997). Developmental stability and fitness: a review. *The American Naturalist*, 149, 916-932.
- Møller, A. P., Soler, M., & Thornhill, R. (1995). Breast asymmetry, sexual selection, and human reproductive success. *Ethology and Sociobiology*, 16, 207-219.

- Møller, A. P., & Thornhill, R. (1998). Bilateral symmetry and sexual selection: a meta-analysis. *The American Naturalist*, 151, 174-192.
- Møller, A. P., & Thornhill, R. (1997). A meta-analysis of the heritability of developmental stability. *Journal of Evolutionary Biology*, 10, 1-16.
- Morris, J. S., Friston, K. J., Buchel, C., Frith, C. D., Young, A. W., Calder, A. J., & Dolan, R. J. (1998). A neuromodulatory role for the human amygdala in processing emotional facial expressions. *Brain*, 121, 47-57.
- Morris, J. S., Frith, C. D., Perrett, D. I., Rowland, D., Young, A. W., Calder, A. J., & Dolan, R. J. (1996). A differential neural response in the human amygdala to fearful and happy facial expressions. *Nature*, 383, 812-815.
- Morris, J. S., Öhman, A., & Dolan, R. J. (1999). A subcortical pathway to the right amygdala mediating 'unseen' fear. *Proceedings of the National Academy of Sciences, USA*, 96, 1680-1685.
- Murray, L., Wheeldon, T., Reid, I., Rowland, D., Burt, D., & Perrett, D. (in submission). Depression and facial expression sensitivity: exploratory studies. *Cognition and Emotion*.
- Murray, L. K., & Perrett, D. I. (in submission). The influence of mood state on sensitivity to different facial expressions. *Journal of Cognitive Neuroscience*.
- Nelson, C. A., Morse, P. A., & Leavitt, L. A. (1979). Recognition of facial expressions by 7-month-old infants. *Child Development*, 50, 1239-1242.
- Nicholls, M. R. N. (1995). Hemispheric asymmetries for temporal resolution: a signal detection analysis of threshold and bias. *Quarterly Journal of Experimental Psychology*, 47A, 291-310.
- Niedenthal, P. M., & Halberstadt, J. (1997). Being happy and seeing "happy": emotional state mediates visual word recognition. *Cognition and Emotion*, 11, 403-432.
- Niedenthal, P. M., Halberstadt, J., Margolin, J., & Innes-Ker, A. (2000). Emotional state and the detection of change in facial expression of emotion. *European Journal of Social Psychology*, 30, 211-222.
- O'Toole, A. J., Deffenbacher, K. A., Valentine, D., & Abdi, H. (1994). Structural aspects of face recognition and the other race effects. *Memory and Cognition*, 22, 208-224.
- O'Toole, A. J., Price, T., Vetter, T., Barrlett, J. C., & Blanz, V. (1999a). 3D shape and 2D surface textures of human faces: the role of "averages" in attractiveness and age. *Image and Vision Computing*, 18, 9-19.

- O'Toole, A. J., Vetter, T., & Blanz, V. (1999b). Three-dimensional shape and two-dimensional surface reflectance contributions to face recognition: an application of three-dimensional morphing. *Vision Research*, 39, 3145-3155.
- O'Toole, A. J., Vetter, T., Troje, N. F., & Bulthoff, H. H. (1997a). Sex classification is better with three-dimensional head structure than with image intensity information. *Perception*, 26, 75-84.
- O'Toole, A. J., Vetter, T., Volz, H., & Salter, E. M. (1997b). Three-dimensional caricatures of human heads: Distinctiveness and the perception of facial age. *Perception*, 26, 719-732.
- Parsons, P. A. (1992). Fluctuating asymmetry: A biological monitor of environmental and genomic stress. *Heredity*, 68, 361-364.
- Pelchat, M. L., Harvey, J. G., Rozin, P., & Jacobs, J. (1983). Quality of acquired responses to tastes in *Rattus norvegicus* depends on type of associated discomfort. *Journal of Comparative Psychology*, 97, 140-153.
- Penn, D. J., & Potts, W. K. (1999). The Evolution of Mating Preferences and Major Histocompatibility Complex Genes. *American Naturalist*, 153, 145-164.
- Penton-Voak, I. S., Jones, B. C., Little, A. C., Baker, S., Tiddeman, B., Burt, D. M., & Perrett, D. I. (2001). Symmetry, sexual dimorphism in facial proportions and male facial attractiveness. *Proceedings of the Royal Society of London Series B-Biological Sciences*, 268, 1617-1623.
- Penton-Voak, I. S., Perrett, D. I., Castles, D. L., Kobayashi, T., Burt, D. M., Murray, L. K., & Minamisawa, R. (1999a). Menstrual cycle alters face preference. *Nature*, 399, 741-742.
- Penton-Voak, I. S., Perrett, D. I., & Pierce, J. (1999b). Computer graphic studies of the role of facial similarity in judgements of attractiveness. *Current Psychology*, 18, 104-117.
- Perrett, D. I., Burt, D. M., Penton-Voak, I. S., Lee, K. J., Rowland, D. A., & Edwards, R. (1999). Symmetry and human facial attractiveness. *Evolution and Human Behavior*, 20, 295-307.
- Perrett, D. I., Lee, K. J., Penton-Voak, I., Rowland, D., Yoshikawa, S., Burt, D. M., Henzi, S. P., Castles, D., & Akamatsu, S. (1998). Effects of sexual dimorphism on facial attractiveness. *Nature*, 394, 884-887.
- Perrett, D. I., May, K. A., & Yoshikawa, S. (1994). Facial shape and judgments of female attractiveness. *Nature*, 368, 239-242.



- Perrett, D. I., Smith, P. A. J., Potter, D. D., Mistlin, A. J., Head, A. S., Milner, A. D., & Jeeves, M. A. (1984). Neurones responsive to faces in the temporal cortex: Studies of functional organization, sensitivity to identity and relation to perception. *Human Neurobiology*, 3, 197-208.
- Persad, S. M., & Polivy, J. (1993). Differences between depressed and nondepressed individuals in the recognition of and response to facial emotional cues. *Journal of Abnormal Psychology*, 102, 358-368.
- Phillips, M. L., Young, A. W., Senior, C., Brammer, M., Andrew, C., Calder, A. J., Bullmore, E. T., Perrett, D. I., Rowland, D., Williams, S. C. R., Gray, J. A., & David, A. S. (1997). A specific neural substrate for perceiving facial expressions of disgust. *Nature*, 389, 495-498.
- Pittenger, J. B., & Shaw, R. E. (1975). Aging faces as viscal-elastic events: implications for a theory of non-rigid shape perception. *Journal of Experimental Psychology: Human Perception and Performance*, 1, 374-382.
- Preston, S. D., & de Waal, F. B. M. (2000). Empathy: its ultimate and proximate bases. *Brain and Behavioural Sciences*, in submission, <http://cogprints.soton.ac.uk/documents/disk0/00/00/10/42/>.
- Previc, F. H. (1991). A general theory concerning the prenatal origins of cerebral lateralization in humans. *Psychological Review*, 98, 299-334.
- Rhodes, G. (1986). Perceptual asymmetries in face recognition. *Brain and Cognition*, 4, 197-218.
- Rhodes, G. (1993). Configural coding, expertise, and the right hemisphere advantage for face processing. *Brain and Cognition*, 22, 19-41.
- Rhodes, G., Brennan, S., & Carey, S. (1987). Identification and ratings of caricatures - implications for mental representations of faces. *Cognitive Psychology*, 19, 473-497.
- Rhodes, G., Halberstadt, J., & Brajkovich, G. (2001a). Generalization of mere exposure effects to averaged composite faces. *Social Cognition*, 19, 57-70.
- Rhodes, G., Hickford, C., & Jeffery, L. (2000). Sex-typicality and attractiveness: Are supermale and superfemale faces super-attractive. *British Journal of Psychology*, 91, 125-140.
- Rhodes, G., & McLean, I. G. (1990). Distinctiveness and expertise effects with homogeneous stimuli - towards a model of configural coding. *Perception*, 19, 773-794.

- Rhodes, G., Proffitt, F., Grady, J. M., & Sumich, A. (1998). Facial symmetry and the perception of beauty. *Psychonomic Bulletin and Review*, 5, 659-669.
- Rhodes, G., & Tremewan, T. (1996). Averageness, exaggeration, and facial attractiveness. *Psychological Science*, 7, 105-110.
- Rhodes, G., Yoshikawa, S., Clark, A., Lee, K., McKay, R., & Akamatsu, S. (2001b). Attractiveness of facial averageness and symmetry in non- Western cultures: In search of biologically based standards of beauty. *Perception*, 30, 611-625.
- Rhodes, G., Zebrowitz, L. A., Clark, A., Kalick, S. M., Hightower, A., & McKay, R. (2001c). Do facial averageness and symmetry signal health? *Evolution and Human Behavior*, 22, 31-46.
- Ritter, J. M., & Langlois, J. H. (1988). The role of physical attractiveness in the observation of adult child interactions - eye of the beholder or behavioral reality. *Developmental Psychology*, 24, 254-263.
- Rizzolatti, G., & Arbib, M. A. (1998). Language within our grasp. *Trends in Neurosciences*, 21, 188-194.
- Rizzolatti, G., Fadiga, L., Matelli, M., Bettinardi, V., Paulesu, E., Perani, D., & Fazio, F. (1996). Localisation of grasp representation in humans by PET - 1; observation vs execution. *Experimental Brain Research*, 111, 246-54.
- Rolls, E. T., & Baylis, G. C. (1986). Size and contrast have only small effects on the responses to faces of neurons in the cortex of the superior temporal sulcus of the monkey. *Experimental Brain Research*, 65, 38-48.
- Rolls, E. T., Toveé, M. J., & Panzeri, S. (1999). The neurophysiology of backward visual masking: Information analysis. *Journal of Cognitive Neuroscience*, 11, 300-311.
- Rosenstein, D., & Oster, H. (1988). Differential facial responses to 4 basic tastes in newborns. *Child Development*, 59, 1555-1568.
- Rowland, D. A. (1998). *Computer graphic control over human face and head appearance, genetic optimisation of perceptual characteristics*. Unpublished Phd, University of St Andrews, St Andrews.
- Rowland, D. A., & Perrett, D. I. (1995). Manipulating facial appearance through shape and color. *IEEE Computer Graphics and Applications*, 15, 70-76.
- Rowland, D. A., Perrett, D. I., Burt, D. M., Lee, K. J., & Akamatsu, S. (1997). *Transforming Facial Images in 2 and 3-D*. Paper presented at the Imagina 97, Monte-Carlo.

- Rubenstein, A. J., Langlois, J. H., & Roggman, L. A. (2001). What makes a face attractive and why: the role of averageness in defining facial beauty. In G. Rhodes & L. Zebrowitz (Eds.), *Facial Attractiveness: Evolutionary, Cognitive and Social Perspectives* (Vol. 1, ). Westport, Connecticut: Ablex.
- Rubenstein, A. J., Lanlgois, J. H., & Kalakanis, L. E. (1999). Infant preferences for attractive faces: a cognitive explanation. *Developmental Psychology*, 35, 848-885.
- Ruppert, J. (1995). A Delaunay refinement algorithm for quality mesh generation. *Journal of Algorithms*, 18, 548-585.
- Russell, K. B. (1997). *Eigenheads for reconstruction*. Unpublished Bachelor of Science in Electrical Engineering and Computer Science, MIT, Cambridge.
- Samuels, C. A., Butterworth, G., Roberts, T., Graupner, L., & Hoyle, G. (1994). Facial aesthetics: babies prefer attractiveness to symmetry. *Perception*, 23, 823-831.
- Scheib, J. E., Gangestad, S. W., & Thornhill, R. (1999). Facial attractiveness, symmetry and cues of good genes. *Proceedings of the Royal Society of London Series B-Biological Sciences*, 266, 1913-1917.
- Schiff, B. B., & Truchon, C. (1993). Effect of unilateral contraction of hand muscles of perceiver biases in the perception of chimaeric and neutral faces. *Neuropsychologia*, 31, 1351-1368.
- Schweinberger, S. R., Burton, A. M., & Kelly, S. W. (1999). Asymmetric dependencies in perceiving identity and emotion: experiments with morphed faces. *Perception and Psychophysics*, 61, 1102-1115.
- Sergent, J. (1982). About face: left-hemisphere involvement in processing physiognomies. *Journal of Experimental Psychology: Human Perception and Performance*, 8, 1-24.
- Simpson, J. A., Gangestad, S. W., Christensen, P. N., & Leck, K. (1999). Fluctuating asymmetry, sociosexuality, and intrasexual competitive tactics. *Journal of Personality and Social Psychology*, 76, 159-172.
- Singh, D. (1993). Adaptive significance of female attractiveness: role of waist-to-hip ratio. *Journal of personality and social psychology*, 65, 293-307.
- Singh, D. (1995). Female health, attractiveness, and desirability for relationships: role of breast asymmetry and waist-to-hip ratio. *Ethology and Sociobiology*, 16, 465-485.

- Skinner, M., & Mullen, B. (1991). Facial asymmetry in emotional expression: A meta-analysis of research. *British Journal of Social Psychology*, 30, 113-124.
- Snodgrass, J., Levy-Berger, G., & Haydon, M. (1985). *Human experimental psychology*. New York: Oxford University Press.
- Snowdon, C., & Boe, C. (2000). *Social communication about adulterated food in cotton-top tamarins*. Paper presented at the 37th Annual Meeting of the Animal Behaviour Society, Morehouse College and Zoo Atlanta.
- Soussignan, R., & Benoist, S. (1996). Children's facial responsiveness to odors: influences of hedonic valence of odor, gender, age, and social presence. *Developmental Psychology*, 32, 367-379.
- Sprenkelmeyer, R., Young, A. W., Calder, A. J., Karnat, A., Lange, H., Homberg, V., Perrett, D. I., & Rowland, D. (1996). Loss of disgust - perception of faces and emotions in Huntington's disease. *Brain*, 119, 1647-1665.
- Steiner, J. E. (1973). The gustofacial response: observation on normal and anencephalic newborn infants. *Symposium on Oral Sensory Perception*, 4, 254-278.
- Steiner, J. E. (1974). Innate discriminative human facial expression to taste and smell stimulation. *Annals of the New York Academy of Sciences*, 237, 229-233.
- Steyvers, M. (1999). Morphing techniques for manipulating face images. *Behavior Research Methods Instruments and Computers*, 31, 359-369.
- Stoddard, J. T. (1886). Composite portraiture. *Science*, 8, 89-91.
- Swaddle, J. P., & Cuthill, I. C. (1995). Asymmetry and human facial attractiveness: symmetry may not always be beautiful. *Proceedings of the Royal Society of London Series B-Biological Sciences*, 261, 111-116.
- Thornhill, R., SW. (1993). Human fluctuating asymmetry and sexual behavior. *Psychological science*, 5, 287-302.
- Thornhill, R., & Gangestad, S. W. (1993). Human facial beauty: averageness, symmetry and parasite resistance. *Human Nature*, 4, 237-269.
- Thornhill, R., & Gangestad, S. W. (1999). Facial attractiveness. *Trends in Cognitive Sciences*, 3, 452-460.
- Thornhill, R., Gangestad, S. W., & Comer, R. (1995). Human female orgasm and mate fluctuating asymmetry. *Animal Behaviour*, 50, 1601-1615.
- Thornhill, R., & Moller, A. P. (1997). Developmental stability, disease and medicine. *Biological Reviews of the Cambridge Philosophical Society*, 72, 497-548.

- Tiddeman, B. (2001). Face Morphing Exhibit. St Andrews: Perception Lab.
- Tiddeman, B., Burt, D., & Perrett, D. (2001a). Computer graphics in facial perception research. *IEEE Computer Graphics and Applications*, 21, 42-50.
- Tiddeman, B., Duffy, N., & Rabey, G. (2001b). A general method for overlap control in image warping. *Computers and Graphics*, 25, 59-66.
- Tiddeman, B., & Perrett, D. (2001). *Moving facial image transformations using Static 2D prototypes*. Paper presented at the 9-th International Conference in Central Europe on Computer Graphics, Plzen, Czech Republic.
- Tovee, M. J., & Cornelissen, P. L. (1999). The mystery of female beauty. *Nature*, 399, 215-216.
- Tovee, M. J., Tasker, K., & Benson, P. J. (2000). Is symmetry a visual cue to attractiveness in the human female body? *Evolution and Human Behavior*, 21, 191-200.
- Trivers, R. (1972). Parental investment and sexual selection. In B. Campbell (Ed.), *Sexual Selection and the Decent of Man* (pp. 136-179). Chicago: Aldine.
- Turk, M., & Pentland, A. (1991). Eigenfaces for recognition. *Journal of Cognitive Neuroscience*, 3, 71--86.
- Vetter, T. (1998). Synthesis of novel views from a single face image. *International Journal of Computer Vision*, 28, 103-116.
- Vetter, T., Jones, M. J., & Poggio, T. (1997). *A bootstrapping algorithm for learning linear models of object classes*. Cambridge, MA: Massachusetts Institute of Technology.
- Walker, E., McGuire, M., & Bettes, B. (1984). Recognition and identification of facial stimuli by schizophrenics and patients with affective-disorders. *British Journal of Clinical Psychology*, 23, 37-44.
- Wallis, G., & Bülthoff, H. H. (2001). Effects of temporal association of recognition memory. *Proceeding of the National Academy of Sciences*, 98, 4800-4804.
- Webster, M. A., & MacLin, O. H. (1999). Figural aftereffects in the perception of faces. *Psychonomic Bulletin and Review*, 6, 647-653.
- Whalen, P. J., Rauch, S. L., Etcoff, N. L., McInerney, S. C., Lee, M. B., & Jenike, M. A. (1998). Masked presentation of emotional facial expressions modulates amygdala activity without explicit knowledge. *Journal of Neuroscience*, 18, 411-418.

- Wolberg, G. (1990). *Digital image warping*. Los Alamitos, CA: IEEE computer society press.
- Wolf, M. E., & Goodale, M. A. (1987). Oral asymmetries during verbal and non-verbal movements of the mouth. *Neuropsychologia*, 25, 375-396.
- Wolff, W. (1933). The experimental study of forms of expression. *Character and Personality*, 2, 168-173.
- Wylie, D. R., & Goodale, M. A. (1988). Left-sided oral asymmetries in spontaneous but not posed smiles. *Neuropsychologia*, 26, 823-832.
- Yamada, H., Chiba, H., Tsuda, K., Maiya, K., & Harashima, H. (1992). A facial image processing system for psychological studies. *Proceedings IEEE International Workshop on Robot Human Communication*, 69, 358-362.
- Yamaguchi, M., Oda, M., & Fukamachi, H. (1995). Does cardioid strain change in real front view face images have an effect on changing the perceived age? *ATR Technical Reports*.
- Yoshikawa, S., & Wataru, S. (2001). Representational momentum in the perception of dynamic facial expressions. *International Journal of Psychology*, 35, 286.
- Young, A., Rowland, D., Calder, A., Etcoff, N., Seth, A., & Perrett, D. (1997). Facial expression megamix: tests of dimensional and category accounts of emotion recognition. *Cognition*, 63, 271-313.
- Young, A. W., Hellawell, D., & Hay, D. C. (1987). Configurational information in face perception. *Perception*, 16, 747-759.
- Zaidel, D. W., Chen, A. C., & German, C. (1995). She is not a beauty even when she smiles: possible evolutionary basis for a relationship between facial attractiveness and hemispheric specialization. *Neuropsychologia*, 33, 649-655.
- Zhang, Z. (2001). *Image-based modeling of objects and human faces*. Paper presented at the SPIE Conference on Videometrics and Optical Methods for 3D Shape Measurement Proceedings, San Jose, CA.
- Zhao, L., & Chubb, C. (2001). The size-tuning of the face-distortion after-effect. *Vision Research*, 41, 2979-2994.



## **Appendix to Chapter 2: Colour caricaturing of facial attractiveness.**

## **Background**

Perrett *et al.* (1994) used averaging to show the overall average face shape is less attractive than the face shape formed by averaging only the most attractive faces, and that the more attractive average face shape can be made more attractive by accentuating the shape differences between it and average. Using an analogous techniques it is possible to caricature average colouration (Rowland & Perrett, 1995). In the experiment presented here we test whether the colouration of faces influences perception facial attractiveness and, if so, whether colour cues present in the more attractive faces might be exaggerated to enhance beauty.

It was hypothesised that colouration cues to attractiveness would show in averages so that the average made from the most attractive faces would be preferred by subjects to the average colouration of the least attractive faces. Furthermore it is expected that averages made from the most attractive faces will be preferred over averages made from a cross-section of the population. Although it is noted that the cross-section is a better representation of the average of the population and therefore might, given a rigid view of Langlois and Roggman's (1990) "Attractive faces are only average" hypothesis, be expected to be found more attractive by subjects.

Perrett *et al.* (1994) found that shape caricaturing to accentuate the traits of the shape average shape formed from the most female faces increased perceived attractiveness. (Although, a similar effect was not found for male faces.) It is expected that colour may be manipulable in a similar fashion and therefore accentuating average colour differences between the most and least attractive faces will have the effect of making a more attractive face image when the accentuation is in the direction of the colouration of the more attractive faces and will have the effect of making a

less attractive face image when the accentuation is in the direction of the less attractive face images.

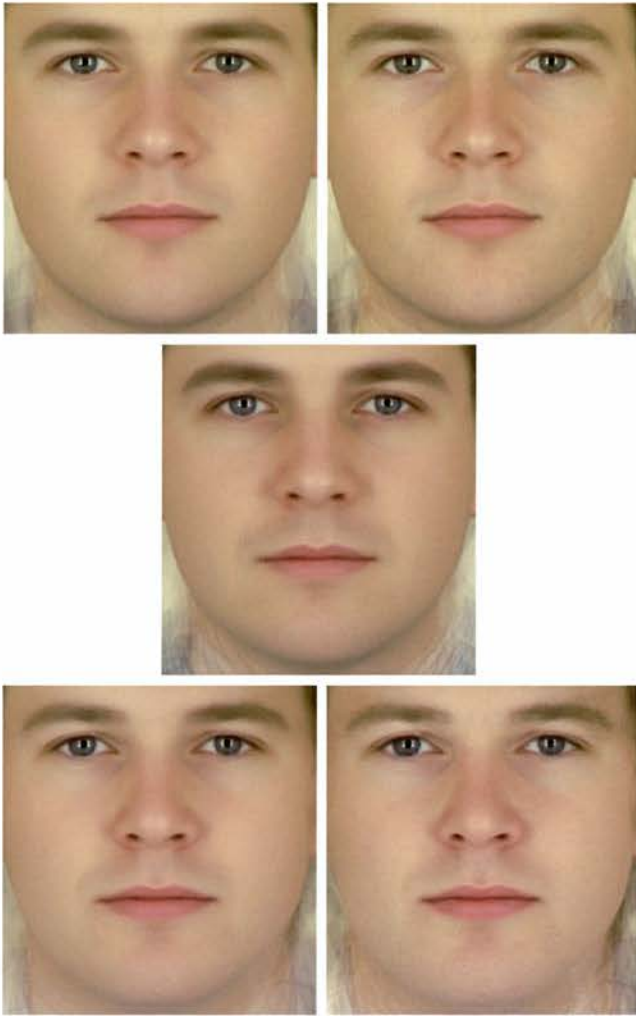
## **Methods**

### **Stimulus manufacture**

15 female faces and 15 male faces (aged 20-29 years as used by Perrett et al., 1994) were rated for attractiveness. The ratings were used to make 4 groups for each sex: a *low* group consisting of the 15 individuals rated as being least attractive, a *high* group consisting of the 15 individuals rated as being most attractive, a *cross-section* group consisting of 15 representing all levels of attractiveness and an *all* group representing all of the faces of that gender. The *cross-section* group was used to make the colouration of the average to keep the number of images used in the blend image constant since the use of greater numbers of face images in a blend tends to lead to greater blurring of the average image.

To manufacture the colour stimuli all face images were then warped into the *all* shape for their gender. For the *high*, *low* and *cross-section* groups superimposing the 15 face images in the group formed a colour average. These colour averages were then used to make 100% caricatures: a *high caricature* by adding the difference between the *low* and the *high* colour average to the *cross-section* colour average and were used to make a *low caricature* by subtracting the same difference from the *cross-section* average (for more details of the image processing techniques see Rowland & Perrett, 1995).

All of the manufactured images were then cropped to 309 by 330 pixels. Example stimuli can be seen in Figures A2.6 and A2.7.



**Figure A2.6: Male stimuli for the colour caricaturing of attractiveness. The male face colour *high* (top left), *high caricature* (top right), *cross-section* (centre), *low* (bottom left) and *low caricature* (bottom right).**

## **Stimulus presentation**

Stimuli were presented in randomly ordered pairs. Subjects were asked to select for each trial the stimulus they preferred.

## **Subjects**

39 subjects (average age 20 years) of whom 24 were female completed the experiment.



Figure A2.7: Female stimuli for the colour caricaturing of attractiveness. The female face colour *high* (top left), *high caricature* (top right), *cross-section* (centre), *low* (bottom left) and *low caricature* (bottom right).

## Results

For each comparison a binomial test was performed. The results of which are shown in Table A2.1. For both male and female stimuli, apart from one exception, subjects had a significant preference in the expected direction i.e. the *high* stimulus was preferred over the *low*, the *low* stimulus was preferred over the *low caricature*, the *high* stimulus was preferred over the *cross* and the *high caricature* was preferred over the *high*. This exception was for the female *high* and *high caricature* for which there was no preference for the *high caricature*.

		comparison	% Preference for 1st stimulus	P value
Stimulus gender	male	<i>high</i> vs <i>low</i>	84%	< 0.001
		<i>high</i> vs <i>cross</i>	79%	< 0.001
		<i>high caricature</i> vs <i>high</i>	74%	< 0.01
		<i>low caricature</i> vs <i>low</i>	13%	< 0.001
	female	<i>high</i> vs <i>low</i>	95%	< 0.001
		<i>high</i> vs <i>cross</i>	87%	< 0.001
		<i>high caricature</i> vs <i>high</i>	63%	0.144
		<i>low caricature</i> vs <i>low</i>	8%	< 0.001

Table A2.1: Binomial comparison of colour stimuli.

## Discussion

Results show that for both male and female faces that the averages made from the most attractive faces are thought more attractive than averages made from less attractive faces. Thus, using image blending colouration cues can be captured. Blends derived from a *cross-section* of the images (and were thus a better representation of the overall population) were preferred less than the blend derived from the most attractive faces. This result would not be predicted from a strict version of Langlois and Roggman's (1990) "Attractive faces are only average" hypothesis. Furthermore, the use of caricaturing in the direction of the *high* attractive blend resulted in male face images that were preferred by subjects over the original *high* attractive average colouration image. Thus manipulation of the colouration of the face images to make them appear less average can result in more attractive faces, a result that is also at odds with Langlois and Roggman's (1990) hypothesis.

Caricaturing produces a slight roughness in the texture of the images as can be seen from Figures A2.6 and A2.7. This may account to the perception of the female *high caricature* not being perceived as more attractive than the female *high* blend by a significant proportion of subjects.



The definition used here of colouration is a procedural description. Each of the images in Figure 2.6 and 2.7 are nominally of the same shape (the average shape calculated from all of the male or female face images used in the experiment). Despite this each image within a figure seems to have a different 3-D shape. This apparent shape is from cues based on the shadow information from which we are able to infer 3-d shape. The procedural nature of the definition of colour used here makes it impossible to determine the reason why the one of the images is preferred over the other, whether it is preferred because of its colouration, texture or 3-d shape (as given but shape from shadow cues).

## **Appendix to Chapter 3.**

The BDI questionnaire (Beck et al., 1961) in the manner presented to students performing Experiment 4 in Chapter 3:

Please read each statement and then choose the statement from each group which best describes the way you have been feeling the past week, including today.

1. ☐ I do not feel sad.  
☐ I feel sad.  
☐ I am sad all the time and I can't snap out of it.  
☐ I am so sad or unhappy that I can't stand it.
2. ☐ I am not particularly discouraged about the future.  
☐ I feel discouraged about the future.  
☐ I feel I have nothing to look forward to.  
☐ I feel that the future is hopeless and that things cannot improve.
3. ☐ I do not feel like a failure  
☐ I feel I have failed more than the average person.  
☐ As I look back on my life, all I can see is a lot of failures  
☐ I feel I am a complete failure as a person.
4. ☐ I get as much satisfaction out of things as I used to.  
☐ I don't enjoy things the way I used to.  
☐ I don't get real satisfaction out of anything anymore  
☐ I am dissatisfied or bored with everything
5. ☐ I don't feel particularly guilty.  
☐ I feel guilty a good part of the time.  
☐ I feel quite guilty most of the time.  
☐ I feel guilty all of the time

6. ☐ I don't feel I am being punished.  
☐ I feel I may be punished.  
☐ I expect to be punished.  
☐ I feel I am being punished.
7. ☐ I don't feel disappointed in myself.  
☐ I am disappointed in myself.  
☐ I am disgusted with myself  
☐ I hate myself.
8. ☐ I don't feel I am any worse than anybody else.  
☐ I am critical of myself for my weaknesses or mistakes  
☐ I blame myself all the time for my faults.  
☐ I blame myself for everything bad that happens
9. ☐ I don't have any thoughts of killing myself.  
☐ I have thoughts of killing myself, but I would not carry them out.  
☐ I would like to kill myself  
☐ I would kill myself if I had the chance.
10. ☐ I don't cry any more than usual.  
☐ I cry more now than I used to.  
☐ I cry all the time now.  
☐ I used to be able to cry, but now I can't cry even though I want to.
11. ☐ I am no more irritated now than I ever am.  
☐ I get annoyed or irritated more easily than I used to.  
☐ I feel irritated all the time now.  
☐ I don't get irritated at all by the things that used to irritate me.

12. ☐ I have not lost interest in other people.
- ☐ I am less interested in other people than I used to be.
- ☐ I have lost most of my interests in other people.
- ☐ I have lost all of my interest in other people.
13. ☐ I make decisions about as well as I ever could.
- ☐ I put off making decisions more than I used to.
- ☐ I have greater difficulty in making decisions than before.
- ☐ I can't make decisions at all anymore.
14. ☐ I don't feel I look any worse that I used to.
- ☐ I am worried that I am looking old or unattractive..
- ☐ I feel that there are permanent changes in my appearance that make me look unattractive.
- ☐ I believe that I look ugly.
15. ☐ I can work about as well as before.
- ☐ It takes an extra effort to get started at doing something.
- ☐ I have to push myself very hard to do anything.
- ☐ I can't do any work at all.
16. ☐ I can sleep as well as usual.
- ☐ I don't sleep as well as I used to.
- ☐ I wake up 1-2 hours earlier than usual and find it hard to get back to sleep.
- ☐ I wake up several hours earlier than I used to and cannot get back to sleep.
17. ☐ I don't get more tired than usual.
- ☐ I get tired more easily than I used to.
- ☐ I get tired from doing almost anything.
- ☐ I am too tired to do anything.

18. ☐ My appetite is no worse than usual.  
☐ My appetite is not as good as it used to be.  
☐ My appetite is much worse now.  
☐ I have no appetite at all anymore.
19. ☐ I haven't lost much weight, if any lately.  
☐ I have lost more than 5 pounds.  
☐ I have lost more than 10 pounds.  
☐ I have lost more than 15 pounds.

I am purposely trying to lose weight by eating less. ☐ Yes ☐ No.

20. ☐ I am no more worried about my health than usual.  
☐ I am worried about physical problems such as aches and pains; or upset stomach or constipation.  
☐ I am very worried about physical problems and it's hard to think of much else.  
☐ I am so worried about my physical problems that I cannot think about anything else.
21. ☐ I have not noticed any recent change in my interest in sex.  
☐ I am less interested in sex than I used to be.  
☐ I am much less interested in sex than I used to be.  
☐ I have lost interest in sex completely.



## **Appendix: Facial affect perception in alcoholics.**

## ***Introduction***

The ability to perceive the emotional state of another person is crucial in everyday social interaction. The face is a complex source of signals and provides important cues for individuals to regulate their own behaviour and interpret the behaviour of others. Subjects who are better at decoding non-verbal cues are more involved in positive social interaction and subjects who are less skilled in decoding facial expression demonstrate less social competence (Feldman, Philippot, & Custrini, 1991; Philippot & Feldman, 1990).

There are a wide range of adverse consequences of alcohol abuse (Duberstein, Conwell, & E.D., 1993; Harford, Grant, & Hasin, 1991; Nixon, Tivis, & Parson, 1992) including severe interpersonal difficulties (disruption of social relationships and family, emotional problems, violence and aggression). It is possible that these interpersonal difficulties are not only a result but also a partial cause of alcoholism. This view is supported by the finding that training to improve communicative skills seems to prevent relapse to alcoholism (Eriksen, Björnstad, & Götestam, 1986; Rohsenow et al., 1991).

Alcohol's association with violent crime has lead to the general belief that it facilitates aggressive behaviour (Edward et al., 1994). This facilitation has been confirmed by studies of intoxicated subjects (Hoaken & Pihl, 2000). Many factors are involved in aggressive behaviour and different models have been used to explain the effect of alcohol on the perception of provocation or threat. For instance, it is known that alcohol has an effect on 5HT, GABA and NMDA receptors and the action on GABA receptors could have an anxiolytic effect reducing the inhibitory impact that fear exercises in response to threat (Gray, 1987).

Alcoholics' aggressive and inappropriate behaviour may also be related to idiosyncrasy in their perception of facial expressions. Oscar-Berman and colleagues (1990) found that Korsakoff and non-Korsakoff alcoholics overestimated the intensity conveyed by full-blown facial expressions. Philippot and colleagues (1999) confirmed this result and extended the findings to moderate, weak and even neutral faces. They found that alcoholics performed worse than controls in a facial expression recognition task and suggested this could be due to a general deficit in visual spatial information processing (Schandler, Clegg, Thomas, & Cohen, 1996). Philippot *et al.* (1999) also found that alcoholic subjects had a systematic bias towards identifying faces as showing anger and contempt; emotions that they note are likely to reflect interpersonal conflict. In this paper we investigate alcoholics' recognition of the different facial expressions but also the impact of attention direction on expression processing.

Attention direction, where an individual appears to be attending to, is known to be highly important in primates, including humans, so much so that we appear to have specialist groups of neurones in the brain 'dedicated' to detecting the attention direction of others. For example, there are neural systems that are selectively activated by different head and gaze directions in the area of the brain called the superior temporal sulcus (Allison, Puce, & McCarthy, 2000; Perrett, Hietanen, Oram, & Benson, 1992). Eye contact is used by many primate species to communicate threat (for example in rhesus monkeys) and in humans gaze serves many functions (Patterson, 1983). Gaze provides information that is used in the evaluation of many different traits including attraction, competence, attentiveness, social skills and dominance. Eye gaze is important in communicating social control and in the

understanding of another person's motives, both of which are crucial for engaging in appropriate interactions (Kleinke, 1986).

Recent research has suggested that the processing of faces displaying expressions of emotion is different depending on whether the expressions are directed towards the subject or directed elsewhere (Yoshikawa & Sato, 2001). The modulating effects of attention direction on attributions made to expressive faces may be important in understanding how alcoholics perceive facial expressions. During eye contact, a perceiver becomes personally involved as the object of another individual's emotion, so specific responses and attributions are appropriate. Eye contact can therefore be viewed as a way of involving others in an emotional interaction. To investigate the modulating effect of attention direction on facial attributions made by alcoholics we developed a new adaptive test using real-time interactive morphing of facial expressions. The test presents animated facial expressions of emotions to the subject either with attention (eye gaze and head) directed towards the subject or with attention directed away from the subject. Each trial in the test is made up of two parts. In the first part subjects identify the facial expression after it has animated from neutral expression to one of 100% intensity. This first part of the test provides a measure of accuracy of expression identification. Subjects then go onto a second part in which they select the minimum amount of expression that they feel is necessary for them to see the expression, thus providing a measure of sensitivity.

Previous work (Philippot et al., 1999) has found that alcoholic subjects are generally less accurate in the identification of expressions and we expected to replicate these findings. Philippot and colleagues also found that alcoholic subjects had a systematic bias in interpreting faces expressing disgust as showing anger or contempt. We expected a similar bias to be present in alcoholic subjects tending to

identify faces looking towards them (as tested by Philippot and colleagues) as showing more anger. As well as using images of individuals that are facing the subject our test uses images in which the individuals making facial expressions direct their attention away from the subject. This enables us to investigate the modulating effect of attention direction on attributions of emotions to faces. We expected the tendency of alcoholic subjects to interpret facial expression as hostile to be less prevalent for faces that are not looking at them.

Our test also provides a measure of sensitivity. This can be used to test a second theory that comes from studies that found a tendency for alcoholics to rate facial expression as being more intense compared to control subjects (Oscar-Berman et al., 1990; Philippot et al., 1999). If facial expressions appear more intense to alcoholic subjects then we might expect that they would select less intense facial expressions when asked to specify the minimum intensity necessary to see the expression. The subjects also completed a control task that was unrelated to facial emotion perception to test if the any changes in sensitivity are specific to the domain of facial expression of emotions.

## ***Method***

### **Participants**

Twenty-five subjects, 13 males and 12 females diagnosed with alcohol dependence according to DSM IV criteria (DSM-IV American Psychiatric Association, 1994) were recruited at a drop-in centre for people with alcohol problems (NOA, Nucleo Operativo Alcolisti) in Milan, Italy. They were undertaking a detoxification program and were not using other psychotropic drugs apart from alcohol. Subjects were tested directly after an interview by their doctor to assess

progress in the detoxification program. All subjects gave informed consent to participate in the study.

Two female subjects were discarded from the data analysis because of dual diagnosis (alcoholism and eating disorders). Patients were matched for age and sex with 23 control subjects free of psychiatric history.

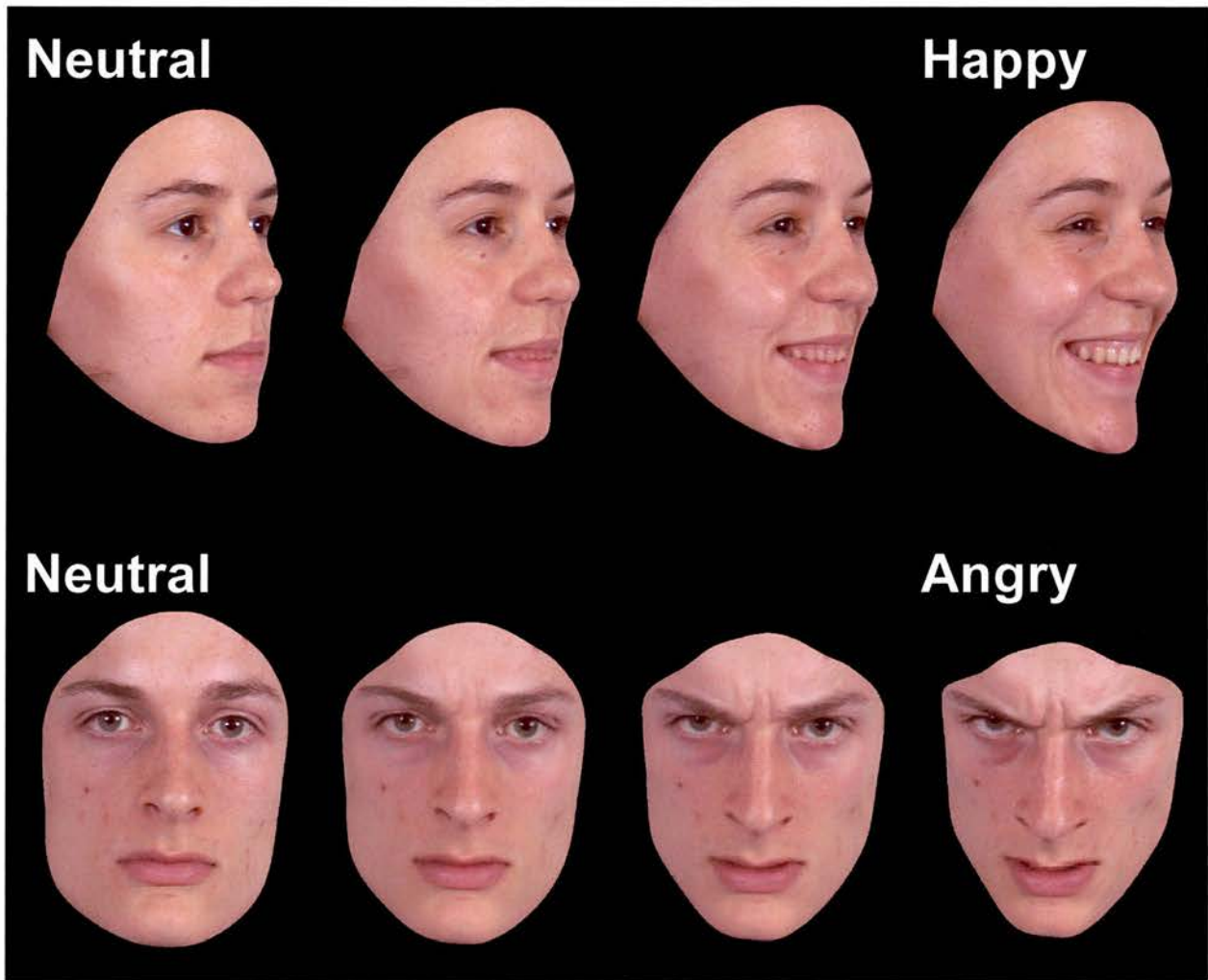
The average age of the alcoholics was 46.7 (SD 10.05) and of the control group was 46.6 (SD 8.78).

## **Materials**

Photographs were taken in frontal and in profile view of 26 individuals asked to make the 6 basic facial expressions (Ekman and Friesen, 1976) of anger, disgust, happiness and sadness. The four individuals (2 males, 2 females) with most recognizable expressions were then selected following a pilot study and used to develop the test. 32 faces (4 identities x 4 expressions x 2 views) were manually delineated using 179 feature-points to define the shape of all the important facial features (Rowland & Perrett, 1995). For example, five delineation points mark the shape and position of the upper eye-lid.

For each of the individuals, 4 morph sequences (Benson & Perrett, 1991) containing 21 images were made representing the change between neutral (0%) and angry, disgusted, happy or sad (100%) expression for both views (32 morph sequences in all). Two examples of sequences are illustrated in Fig. 1. For the practice session a further 4 sequences were made of another individual representing the change from neutral to each of the expressions while facing the camera.





**Fig. 1. Example frames from two expression sequences (0%, 33%, 66%, 100%). The upper sequence is facing-away and the lower facing-towards.**

A ‘control task’ involving the perception of face gender was developed to investigate the specificity of the results obtained from the first test. The stimuli presented were manufactured from 8 pairs of male and female computer composites (Rowland & Perrett, 1995). Each pair of male and female faces was used to form a continuum of 21 images ranging from masculine through androgynous (0%) to feminine. To make the gender of the ends of the continua more visible, the average male and female faces were caricatured in shape and colour away from one another. The composites were as follows: actor (composites of male and female famous actors caricatured to 125%, used in the practice trials), models (composites of male and female models collected from magazines caricatured to 125%), young adults

(composites of males and females aged between 20 and 24 years caricatured to 150%, Burt & Perrett, 1995), old adults (composites of males and females aged between 20 and 24 years caricatured to 150%, Burt & Perrett, 1995), and 3 pairs of composites made from different sets of students' pictures taken under differing photographic conditions, students 1, 2 and 3 each caricatured to 150% (Penton Voak et al., 1999). All composites contained at least 15 individual images.

The Beck Depression Inventory (BDI 21 items questionnaire, Beck, Ward, Mendelson, Mock, & Erbaugh, 1961) was used to assess depressive symptomatology.

## **Procedure**

Subjects were informed about purposes and methods of the research and gave informed consent. Each subject then completed the BDI questionnaire and was given instructions for the computer tasks. Following the 4 practice trials, the 32 randomly ordered trials of the test were presented. Each trial started with a neutral face that animated (morphed) to a full intensity (100%) expression. Subjects had to signal which expression was being produced by clicking on one of four text-labelled icons (angry, disgusted, happy or sad). The face image being displayed then reverted to the neutral (0%) expression and the subject was requested to find the point at which they were first able to see the expression displayed on the face image by using the cursor keys to move back and forth through the animation sequence.

Subjects finished the experiment by performing the control task, consisting of 2 practice trials followed by 14 randomly ordered experimental trials. Subjects saw each stimulus twice: once starting with a highly masculine face, they were asked to detect the position at which the face first looked feminine, and then, starting with a highly feminine face, subjects were asked to detect when the face first looked masculine.

The whole procedure took up to 60 minutes depending on the subject.

## **Results and discussion**

### **BDI score**

The average BDI score of alcoholic subjects was 14.8 (SD = 10.88, range 1-37) compared to 3.3 of the control group (SD = 3.32, range 0 – 9). An independent samples t-test (unequal variance) showed a significant difference between the two groups for BDI scores [ $t(44) = -4.69$ ,  $p < 0.001$ ]. Alcoholic females were significantly more depressed than alcoholic males [22.8 (sd=11.23) compared to 8.6 (sd=5.24) respectively;  $t(21) = -4.03$ ,  $p < 0.001$ ].

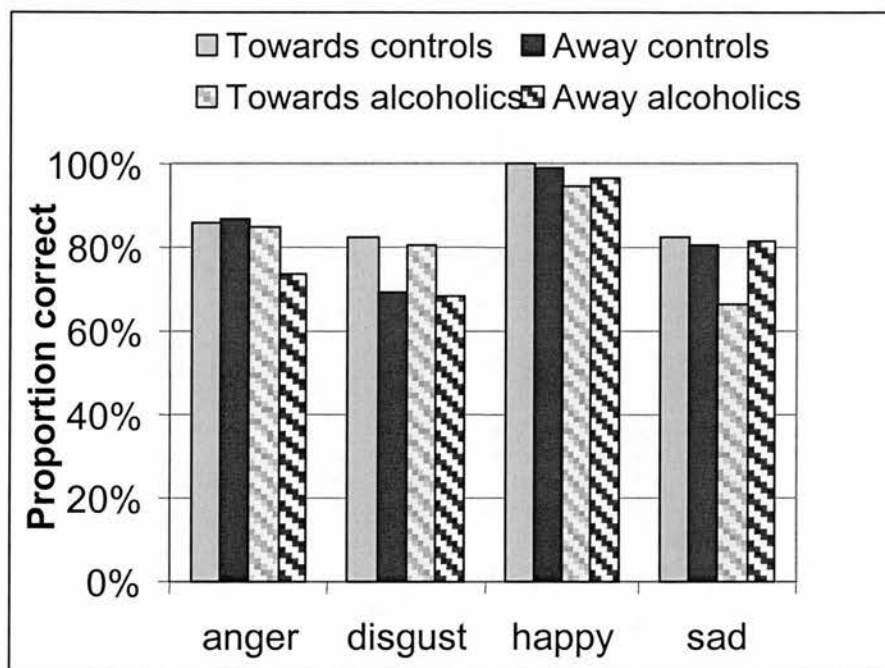
### **Accuracy and perceptual-sensitivity**

Two different indices were calculated: accuracy and perceptual-sensitivity. We define ‘accuracy’ as the percentage of correct identifications for each expression. Perceptual-sensitivity was defined as the average amount of expression intensity that the subjects chose in order to be able to see the expression in the face. Sensitivity was calculated only from trials on which expression was correctly recognised.

### **Recognition accuracy**

An initial analysis of numbers of false positive errors using a 4x2x2x2 mixed-design ANOVA (Analysis of Variance) with Expression (four levels: angry, disgusted, happy and sad) and View (two levels: facing-towards and facing-away) as within-subject factors and Group (two levels: alcoholic and control) and Sex (two levels: male and female) as between-subjects factors revealed only a significant main effect of Expression. Since there were no other effects of false positives (other than

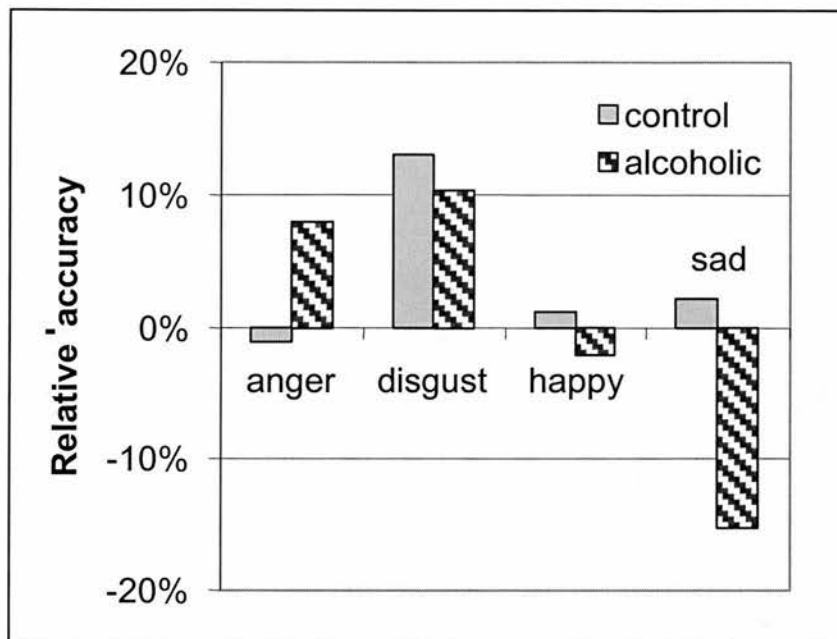
the main effect on Expression) we dismiss effects of false positives on accuracy and proceed to analyse accuracy.



**Fig. 2.** Proportion of expressions labelled correctly for each expression, both views (facing-towards, facing-away), for alcoholic and control subjects.

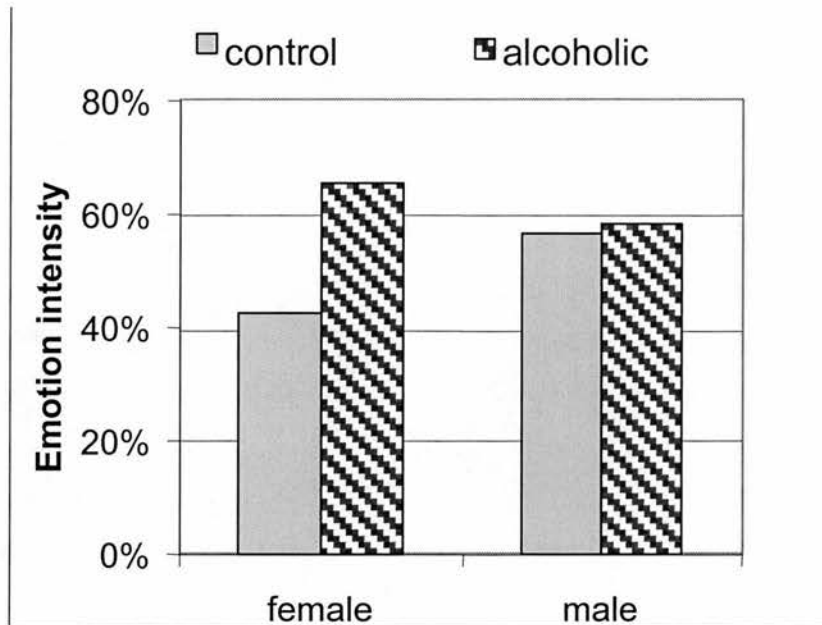
To investigate whether alcoholics perceive facial expressions differently depending upon face orientation, a 4x2x2x2 mixed-design ANOVA was conducted on accuracy scores with Expression (four levels: angry, disgusted, happy and sad) and View (two levels: facing-towards and facing-away) as within-subject factors and Group (two levels: alcoholic and control) and Sex (two levels: male and female) as between-subjects factors. We used the Greenhouse-Geisser correction throughout our analysis as the within-subject factor (Expression) had more than two levels. As illustrated in Fig. 2 the analysis showed:

- a significant main effect of Expression [ $F(2.52, 106.11) = 32.01, p < 0.001$ ]
- a significant main effect of Group [ $F(1, 42) = 4.95, p < 0.05$ ]
- a significant Expression X View interaction [ $F(2.42, 101.74) = 5.23, p < 0.01$ ]
- a significant Expression X View X Group interaction [ $F(2.42, 101.74) = 2.84, p < 0.05$ ]



**Fig. 3. A comparison of alcoholic and control subjects' Relative-accuracy (percentage of correct answers facing-towards minus percentage of correct answers facing-away) for the 4 expressions. A high Relative-accuracy means better recognition of the facing stimuli than the stimuli facing away.**

The main effect of Expression and the Expression X View interaction was due to particular expressions and images (e.g. happy) being easier to recognise than others. The main effect of Group was due to alcoholic subjects being less accurate at identifying expressions than non-alcoholic subjects. To understand the interaction between Expression, View and Group, LSD Post-Hoc tests were computed to compare the accuracy of alcoholic and control subjects for each of the expressions, for each view on Relative-accuracy. Relative-accuracy was calculated as accuracy for facing-towards stimuli minus accuracy for facing-away stimuli. There was a significant difference between alcoholic and control subjects for the expression sad [ $p = 0.009$ ] but not for any of the other expressions: angry [ $p = 0.07$ ], disgusted [ $p = 0.87$ ] or happy [ $p = 0.62$ ] (see Fig. 3 for plot). We interpret this difference as reflecting alcoholic (compared to controls) subjects' bias to incorrectly identify the expressions of faces showing sad (but not other) expressions when directed towards them (rather than away from them). Alcoholic subjects misidentified these sad faces, facing towards them, as being angry or disgusted.



**Fig. 4. The interaction between Group and Sex for Perceived-sensitivity.**

### **Perceptual-sensitivity**

After the trials in which subjects misidentified the expression were discarded from the sensitivity analysis a 4x2x2x2 mixed-design ANOVA was conducted with Expression (four levels: angry, disgusted, happy and sad) and View (two levels: facing-towards and facing-away) as within-subject factors and Group (two levels: alcoholic and control) and Sex (two levels: male and female) as between-subjects factors. The analysis showed (see Fig.4):

- a significant main effect of Expression [ $F(2.76, 115.772) = 33.43$ ,  $p < 0.001$ ]
- a significant main effect of Group [ $F(1, 42) = 9.02$ ,  $p < 0.01$ ]
- a significant Group X Sex interaction [ $F(1, 42) = 6.47$ ,  $p < 0.01$ ]

The main effect of Expression was due to some expressions being easier to perceive than others at lower intensities (e.g. happy). The main effect of Group was caused by alcoholic subjects choosing to add a greater intensity of expression for the expression to be seen as present. This main effect of Group is qualified by the relationship between Group and Sex that is plotted in Fig. 4. In order to understand



the relationship between Group, Sex and Perceptual-sensitivity, Post-Hoc tests were computed. The female controls required a less intensity of expression to be present than female alcoholics [LSD;  $p = 0.001$ ], male alcoholics [ $p = 0.018$ ] and male controls [ $p = 0.039$ ].

### **The control test**

A mixed design ANOVA 2x2x2 with Stimulus-gender (two levels: male to female and female to-male) as a within-subject factor and Group (two levels: alcoholic and control) and Sex (two levels: male and female) as between-subjects factors was computed in order to compare the average intensity that alcoholics and controls need to identify female and male faces. There was a main effect of Stimulus-gender [ $F(1,42) = 7.9$ ,  $p < 0.01$ ] but no significant effect of Group [ $F(1,42) = 1.7$ ,  $p < 0.2$ ] or Sex [ $F(1,42) = 1.1$ ,  $p < 0.3$ ] and no significant interactions indicating that there was no difference between alcoholics and control subjects in performing this task.

### **Discussion**

Our results show that alcoholic subjects made more errors in recognising expressions generally, and had a specific deficit in recognising sad faces directed towards them. Alcoholic subjects tended to interpret sad faces directed at them as expressions of anger or disgust. In the second part of the test, subjects chose the minimum intensity of expression needed to see an expression as being present. We found that female alcoholic subjects needed a greater intensity of expression than female controls. The last part of the study was a control task that did not involve the perception of facial expression. Instead subjects chose the minimum amount of femininity or masculinity needed to identify a face image as male or female. We found that alcoholic subjects did not differ from control subjects on this task.

Consistent with previous studies (Philippot et al., 1999), we found that alcoholic subjects were less accurate at identifying the expressions portrayed by emotional faces. Making more errors in decoding facial expressions might cause more misunderstandings of others' intentions and a greater number of inappropriate reactions which could lead to increased social difficulties and conflicts.

Philippot and colleagues (1999) found that alcoholic subjects have a systematic bias to interpret faces expressing disgust as showing anger or contempt, two emotions that they note are related to interpersonal conflict. In our study alcoholic subjects had a tendency to interpret sad faces that were attending towards them as being hostile (angry and disgusted). This effect was not present in faces that were not looking at the subject. Bias in attributing sad expressions directed towards alcoholics might reflect a hostile attributional bias (Dodge et al., 1990) that can be defined as the tendency to feel more threatened by other individuals and consequently more hostile towards them. All the alcoholics' subjects in our sample have at least one alcoholic parent. This attributional bias could be due to an aggressive response pattern learned early in life (possibly from the alcoholic parents) and might be reactivated every time a threat is perceived. Individuals looking at you may be perceived as more threatening than individuals looking away. Alcoholics, a group that are well known to be hostile (Handelsman et al., 2000) may be more sensitive to feelings of threat from faces looking at them, and so be more likely to interpret facial expressions as reflecting hostile expressions (angry or disgusted) than as sadness.

Another connected interpretation is related to the feeling of low self-esteem that is known to be present in alcoholics (Neeliyara et al., 1988). Alcoholic subjects may be biased to interpreting other people's facial expressions in a manner that reduces their self-esteem. When faced with another individual who has an expression

of sadness on their face most individuals attribute feeling of sadness to the person. Alcoholics on the other hand, may tend to attribute negative ideas to the other individual: that they are angry or disgusted with them. This might be expected to make the alcoholic react angrily and feel more inadequate and compound their low self-esteem.

We found that when asked to select the minimum amount of intensity that is needed to perceive one expression as being present alcoholic subjects select more intensity than control subjects. This finding appears at first glance to be at odds with previous work (Oscar-Berman et al., 1990; Philippot et al., 1999) which found that alcoholics rate facial expression as being more intense than control subjects. We interpret the difference in results as reflecting different methods. Alcoholics are more impulsive (Ketzenberger and Forrest, 2000) and their more intense ratings of facial expressions may relate to their greater impulsivity. The results presented here suggest that alcoholics need a greater intensity of expression to be present and so are less sensitive to facial expressions but are, on the basis of previous studies (Oscar-Berman et al., 1990; Philippot et al., 1999), more reactive to the presence of an expression once they perceive it.

Previous studies have found that female subjects are more sensitive in their perception of expressions than male subjects (Hall et al., 1999). In line with this we found that female control subjects chose lower intensities of expression than other groups.

A more novel result relates to the difference between the female alcoholics and female controls. Female alcoholics chose more intensity in order to see facial expressions. This is possibly due to a lack of expression perception sensitivity in female alcoholics (since the difference in sensitivity between male controls and male

alcoholics was much smaller, and non-significant). This difference in how female and male alcoholics react compared to female and male controls may relate to the female alcoholics higher levels of depression, or be due to any one of the many other differences in alcoholism between the sexes. There are many distinctions in the pattern and consequences of alcohol abuse between male and female. Women tend to start to abuse alcohol later than men, but the time between the beginning of problematic drinking and physical dependence is shorter, health risks are higher and alcohol-related mortality three times greater (Furlan and Picci, 1990; Wilsnack and Beckman, 1984). Women also start problem drinking for different reasons, often because of having to face a crisis linked to their physical well-being or position within their family (Mendelson and Mello, 1992). Female alcoholics tend to hide their drinking and drink alone, possibly because of cultural taboos that are widespread in the western societies (McConville, 1983). Alcohol also has different effects on men and women. Nolem-Hoeksema and colleagues (1999) found that the tendency to ruminate and to drink alcohol is connected, but while alcohol tends to temporarily alleviate ruminations in men, it tends to cause more distress in women.

Alcoholics did not differ from control subjects in the amount of masculinity or femininity they chose to add into the images in order to see them as either male or female in the control task. Therefore we did not find any evidence that alcoholic subjects are poorer in perceiving gender from faces than control subjects. The lack of effect of alcoholism on the control task may suggest that the problems in face processing are not generalised, but they may be more circumscribed and more restricted to particular interpersonal aspects of social cognition.

The present study is the first to investigate the importance of attention direction in modulating alcoholics' attributions of hostility in others. We found that

alcoholics tend to make hostile attributions to individuals looking at them. This may be of importance in understanding the interpersonal behaviour of alcoholics and triggers to the aggression that presents a problem not only for them but also for their carers, family and society in general. More generally this study demonstrates the possible importance of attention direction when investigating psychopathologies.

## References for appendixes

- Allison, T., Puce, A., & McCarthy, G. (2000). Social perception from visual cues: role of the STS region. *Trends in Cognitive Sciences*, 4, 267-278.
- Beck, A. T., Ward, C., Mendelson, M., Mock, J., & Erbaugh, J. (1961). An inventory for measuring depression. *Archives of General Psychiatry*, 4, 53-63.
- Benson, P. J., & Perrett, D. I. (1991). Synthesising continuous-tone caricatures. *Image and Vision Computing*, 9, 123-129.
- Burt, D. M., & Perrett, D. I. (1995). Perception of age in adult Caucasian male faces - computer graphic manipulation of shape and color information. *Proceedings of the Royal Society B*, 259, 137-143.
- Duberstein, P. R., Conwell, Y., & E.D., C. (1993). Interpersonal stressors, substance abuse, and suicide. *Journal of Nervous Mental Disorders*, 181, 80-85.
- Edward, G., Anderson, P., Babor, T. F., Casswell, S., Ferrence, R., Giesbrecht, N., Godfrey, C., Holder, H. D., Lemmens, P., Makela, K., Midanik, L. K., Norstrom, T., Osterberg, E., Romelsjo, A., Room, R., Simpura, J., & Skog, O. J. (1994). *Alcohol policy and the public good*. New York: Oxford University Press.
- Eriksen, L., Bjørnstad, S., & Götestam, K. G. (1986). Social skills training in groups for alcoholics: One year treatment outcome for groups and individuals. *Addictive Behavior*, 11, 309-329.
- Feldman, R. S., Philippot, P., & Custrini, R. J. (1991). Social competence and nonverbal behavior. In R. S. Feldman & R. B. (Eds.), *Fundamentals of Nonverbal Behavior* (pp. 329-350). New York: Cambridge University Press.
- Gray, J. A. (1987). *The Psychology of Fear and Stress*. Cambridge: Cambridge University Press.
- Harford, T. C., Grant, B. F., & Hasin, D. S. (1991). Effect of average daily consumption and frequency of intoxication on the occurrence of dependence symptoms and alcohol-related problems. In W. B. Clark & M. E. Hilton (Eds.), *Alcohol in America: drinking practises and problems*. (pp. 212-37). Albany: State University of New York Press.
- Hoaken, P. N. S., & Pihl, R. O. (2000). The effects of alcohol intoxication on aggressive responses in men and women. *Alcohol and Alcoholism*, 35, 471-477.



- Kleinke, C. L. (1986). Gaze and Eye Contact - a Research Review. *Psychological Bulletin*, 100, 78-100.
- Langlois, J. H., & Roggman, L. A. (1990). Attractive faces are only average. *Psychological Science*, 1, 115-121.
- Nixon, S. J., Tivis, R., & Parson, O. A. (1992). Interpersonal problem-solving in male and female alcoholics. *Alcoholism: Clinical and Experimental Research*, 16, 684-687.
- Oscar-Berman, M., Hancock, M., Mildwordf, B., Hutner, N., & Altman-Weber, D. (1990). Emotional perception and memory in alcoholism and aging. *Alcoholism: Clinical and Experimental Research*, 14, 384-393.
- Patterson, M. (1983). *Nonverbal behaviour: A functional perspective*. New York: Springer.
- Perrett, D. I., Hietanen, J. K., Oram, M. W., & Benson, P. J. (1992). Organization and functions of cells responsive to faces in the temporal cortex. *Philosophical Transactions of the Royal Society of London, Series B.*, 335, 23-30.
- Perrett, D. I., May, K. A., & Yoshikawa, S. (1994). Facial shape and judgments of female attractiveness. *Nature*, 368, 239-242.
- Philippot, P., & Feldman, R. S. (1990). Age and social competence in preschoolers' decoding of facial expression. *British Journal of Social Psychology*, 29, 43-54.
- Philippot, P., Kornreich, C., Blairy, S., Baert, I., Dulk, A. D., Le Bon, O., Streel, E., Hess, U., Pelc, I., & Verbanck, P. (1999). Alcoholics' deficit in the decoding of emotional facial expression. *Alcoholism: Clinical and Experimental Research*, 23, 1031-1038.
- Rohsenow, D. J., Monti, P. M., Binkoff, J. A., Liepman, M. R., Nirenberg, T. D., & D.B., A. (1991). Patient treatment matching for alcoholic men in communication skills versus cognitive-behavioral mood management training. *Addictive Behavior*, 16, 63-69.
- Rowland, D. A., & Perrett, D. I. (1995). Manipulating facial appearance through shape and color. *IEEE Computer Graphics and Applications*, 15, 70-76.
- Schandler, S. L., Clegg, A. D., Thomas, C. S., & Cohen, M. J. (1996). Visuospatial processing in intoxicated, recently detoxified, and long-term abstinent alcoholics. *Journal of Substance Abuse*, 8, 321-333.